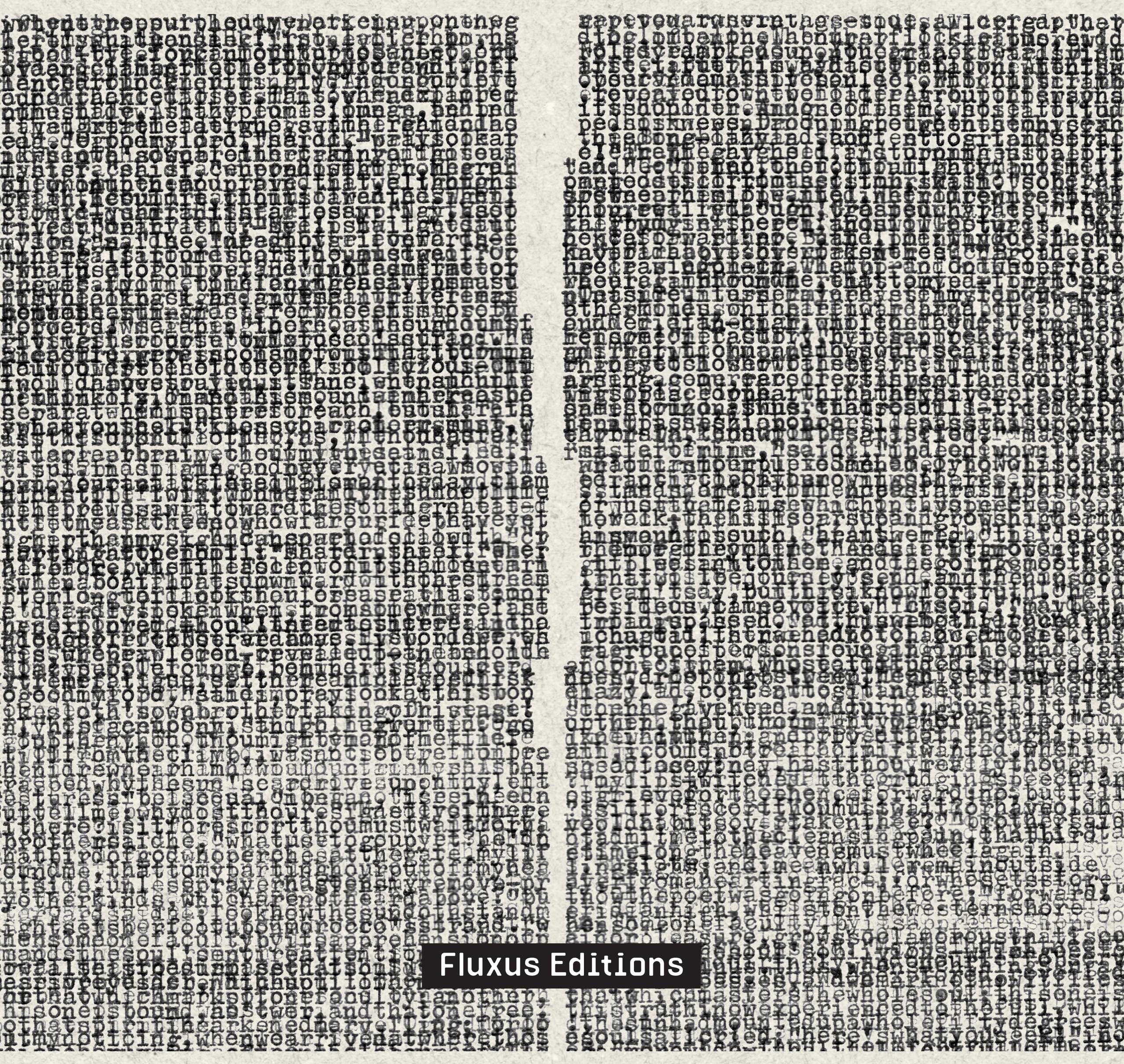
GRAPHOLINGUISTICS AND ITS APPLICATIONS

# The Nature of Writing

## A Theory of Grapholinguistics

Dimitrios Meletis



Fluxus Editions

Grapholinguistics and Its Applications 3

Dimitrios Meletis

# The Nature of Writing

## A Theory of Grapholinguistics

Fluxus Editions

Dimitrios Meletis. 2020. *The Nature of Writing. A Theory of Grapholinguistics.* (Grapholinguistics and Its Applications, Vol. 3). Brest: Fluxus Editions.

This title can be downloaded at:
http://fluxus-editions.fr/gla3.php

ISBN: 978-2-9570549-2-3
e-ISBN: 978-2-9570549-3-0
ISSN: 2681-8566
e-ISSN: 2534-5192
DOI: https://doi.org/10.36824/2020-meletis

Cover illustration: *Purgatory: Canto VII – The Rule of the Mountain* from *A Typographic Dante* (2008) by Barrie Tullett (also displayed in Barrie Tullett, *Typewriter Art: A Modern Anthology*, London: Laurence King Publishing, 2014, p. 167). With kind permission by Barrie Tullett.

The text is taken from *Dante. The Divine Comedy*, translated by Dorothy L. Sayers, Harmondsworth-Middlesex: The Penguin Classics, 1949.

On the lower part of the illustration, one can read the concluding verses of the Canto:

> But now the poet was going on before;
> “Forward!” said he; “look how the sun doth stand
> Meridian-high, while on the Western shore
>
> Night sets her foot upon Morocco’s strand.”

Cover design: Atelier Fluxus Virus
Typesetting: Dimitrios Meletis
Main fonts: William Pro by Typotheque Type Foundry, Computer Modern Typewriter by Donald E. Knuth, Source Han Serif by Adobe Systems, Amiri by Khaled Hosny, Univers by Linotype
Typesetting tools: XƎLATEX, Adobe InDesign.

Fluxus Editions
38 rue Émile Zola
29200 Brest, France
www.fluxus-editions.fr

Dépôt légal : novembre 2020 κηικ

*to my mother, to Lukas, and to Petra*

# Table of contents

# Acknowledgments

First and foremost, I would like to express gratitude to my supervisor Bernhard Hurch – for his support, encouragement, trust, and for sharing his knowledge with me. It is a great feeling to have someone you admire believe in what you do, and it has greatly shaped this finished product. I would also like to thank my second supervisor Christa Dürscheid, for sparing absolutely no efforts to support me and my research, and for being the best mentor I could ever hope for. I am also profoundly grateful to Petra Hödl, my colleague and close friend who has undoubtedly been my biggest supporter over the last couple of years – be it through endless discussions about graphemes, naturalness, or more general linguistic questions, in which she frequently offered me thought-provoking impulses, or many welcome diversions (such as kidnapping a sofa from a vacant office in our department in broad daylight). Sharing an office with her made the process of writing this book not just bearable but memorable and fun. Petra, you are brilliant, and I will be forever in your debt. Many thanks also to Andi Gredig, a new colleague and friend, for our great and helpful conversations about academia and life; I truly treasure them!

For discussions on grapholinguistic topics and feedback on my work, I want to thank Terry Joyce, Peter T. Daniels, Konrad Ehlich, Wolfgang U. Dressler, Florian Busch, Jannis Androutsopoulos, Jürgen Spitzmüller, Daniel Harbour, and Amalia Gnanadesikan, as well as everyone who attended the 10th, 11th, and 12th workshops of the *Association of Written Language and Literacy*, the eikones summer school *Iconicity in writing: practices and constraints* (Basel, September 2016), and the 44th and 46th *LACUS* conferences (Linguistic Association of Canada and the United States). I also want to take this opportunity to thank all the anonymous reviewers who – in different contexts – have given me invaluable feedback. For allowing me to contribute to a great grapholinguistic project and teaching me valuable things along the way, I want to thank Martin Neef. I also thank Kristian Berg, Nanna Fuhrhop, and Karsten Schmidt for sending me unpublished work. For granting me fascinating glimpses into their native languages and writing systems, I am grateful to my teachers Hongling Yang, Amro Elsaidi, Rojana Gottwald, and Amnuaporn Oswald. Furthermore, I benefited greatly from a supportive and collaborative work environment, and I want to thank all of my former colleagues at the University of Graz, especially Veronika Mattes. Dagmar Gramshammer-Hohl has not ceased to support me since my first days as a student, and I will always be grateful to her. For invaluable help with the writing process and important advice about academia in general, I am thankful to Doris Pany-Habsa. I also want to express my gratitude to the Austrian Academy of Sciences for awarding me a fellowship.

For believing in my work and giving me the chance to publish this book in the series *Grapholinguistics and Its Applications* (which is open access, something that was important to me), I am very grateful to Yannis Haralambous. Grapholinguistics is a young and evolving discipline, and his efforts – including this book series as well as the */gʁafematik/* conferences – infuse much-needed life into it, which I think is great. I thank him also for allowing me to publish the book in my design, which (especially as a typography-savy grapholinguist) I absolutely needed to do to conclude this years-long process.

Writing a book like this would be an impossible undertaking without personal support. I owe everything I have and am to my mother, my role model and the strongest person I know, who gave me the chance to pursue my dreams and opened doors for me that remained closed to herself. It was not (financially) self-evident that I could attend university and the fact that I made it to this point is only because of her and her sacrifices. Danke aus tiefstem Herzen, Mama! My partner Lukas is a source of unwavering support. His kindness and belief in me keep me going, and for that, I am thankful beyond words. Sharing our lives fills me with joy. My family and friends always encourage me, even if what I am doing is alien to them. I do not take their support for granted and appreciate it beyond words. Thank you for being a part of my life and enriching it tremendously. I am also grateful to my father for evolving and stepping up – εὐχαριστῶ. Finally, many thanks go to the people on Twitter who motivate me immensely by showing interest in my work.

Last but not least, I want to thank my cats. (Yes, I am serious.) At several stages of my writing process, they selflessly listened to many of my grapholinguistic ramblings (and could not even complain to anyone about it). If certain studies* are to be believed, they have elongated my life by preventing cardiovascular diseases, which are very likely a common side-effect of an academic lifestyle. Thank you for purring.

* Studies such as: Qureshi, Adnan I. et al. (2009): Cat ownership and the risk of fatal cardiovascular diseases. Results from the second national health and nutrition examination study mortality follow-up study. *Journal of Vascular and Interventional Neurology* 2.1: 132–135.

# I Introduction

> […] writing is a system in its own right […] (Vachek 1989: 7)
>
> There are many things that linguists should know about writing: more, in any event, than can be suggested in passing. (Coulmas 1989: 267)

Writing represents language by graphic means. At the same time, it is its own semiotic system. Florian Coulmas (1989: 3) calls it "the single most important sign system ever invented on our planet". Indeed, nowadays, literate cultures are characterized by the ubiquity of writing. While its status as a cultural technique varies across cultures, it appears to differ only in *how* writing is important and not in the fact *that* it is central. With the advent of modern technologies and new forms of communication – many of which are written – writing is gradually gaining even more currency. The fact that written communication unquestionably serves a multitude of registers, contexts, and functions highlights that it has evolved from a resource of the elite and the central cornerstone of high culture to an everyday instrument, echoing the gradual democratization of writing. Academic communication and dissemination of knowledge, too, depend largely on writing, apart from talks at conferences (which, however, are mostly based on or integrate written texts, e.g. in the form of slides or handouts, too) and, to some degree, personal face-to-face communication – contexts in which speech still prevails. Linguistics is no exception to this. Linguistic examples are virtually always rendered in written form. Transcriptions, historical reconstructions – everything in linguistics is written down at some point, except for corpora of spoken language, whose data, however, are also often transcribed. Writing constitutes the very tool that linguistics has always relied on. Yet, ironically, as an object of linguistic study, writing had a pretty rough and delayed start, so to speak – or write.

This is not the place to lament that, following central linguists such as Hermann Paul, Ferdinand de Saussure, Leonard Bloomfield, and other like-minded thinkers (with respect to matters of writing), speech was recognized as the only valuable subject of linguistics and writing was neglected,[1] as this is done abun-

---

1 These linguists are usually listed in the context of the neglect of writing as an object of linguistic study. However, as Glück (1987: Chapter 3) argues, dismissing them as ignorant of writing is shortsighted, as a more elaborate treatment of their works reveals that their opinion of writing was not as one-dimensional as it is frequently depicted.

dantly in a great number of other writing-related publications (cf. Dürscheid 2016: 13–19; Wehde 2000: 43–48; Coulmas 1981: 21–56). In fact, eventually, writing was, if only hesitantly, accepted into the canon of linguistic research interests. As Baroni (2016: 291) observes, “[m]ost linguists, when dealing with graphemics, written language, writing systems and orthography, feel the need to justify themselves. It is about time to change this attitude and to stop feeling guilty about treating graphemics as part of linguistics”. Such a change of attitude has slowly arrived: during the last decades, the investigation of various facets of writing has become gradually more accepted. This, crucially, does not entail that the leading theoretical paradigms in linguistics have ceased to neglect writing. Neither generativism nor, to a large degree, American structuralism or even functionalist enterprises such as usage-based approaches have seriously compelled themselves to include writing in their theories, giving the impression that writing is interpreted as a surplus that need not be included in theories of language. Congruously, while research on writing frequently integrates findings from other linguistic subfields such as phonology and morphology, the same does not apply vice versa – at least not willingly and overtly, given that an implicit *written language bias* (Linell 1982), alternatively known as *scriptist bias* or *scripticism* (Harris 1980; Ágel 2003; Hennig 2006), has been identified. These terms refer to the bizarre situation that writing was (and often still is) dismissed as a secondary and at times unworthy object of study in linguistics despite the fact that all the while, linguistic findings have been based largely on writing. Ágel (2003: 8) demonstrates this using an example of structural/syntactic ambiguity in generativism. A sentence such as “Flying planes can be dangerous”, he claims, is ambiguous only for a reader but not for a hearer, who can potentially resolve the ambiguity by means of prosodic features when the sentence is spoken. This highlights that paradigms such as generativism – with their claims of a modality-indifferent competence – or structuralism – with its (initial) depreciation of writing – have, underneath their explicit claims, actually relied on writing from the very start.

All of this strongly suggests that an investigation of writing is not optional. As a subject, it cannot be put aside for later or placed at the bottom of the list of linguistic priorities. Indeed, the treatment of writing – its features and its categories – is necessary for and, as argued above, actually already inherent in many ‘traditionally’ linguistic analyses. The concepts *phoneme*, *word*, and *sentence*, for instance, are heavily influenced if not constituted by writing. In this vein, some scholars go as far as claiming that the phoneme is epiphenomenal in that its existence is made possible only by the segmentality of alphabetic writing (cf. Faber 1992; Aronoff 1992). In other words, segmental thinking and segmental units of language may only be seizable for us *because of* writing. The strongest version of this view claims that writing does not merely make visible pre-existing units of language but that their visualization in the written modality actually creates these units – in other words, that they are products of graphic relativity (cf. Bugarski 1993; Davidson 2019). Even if this claim were only partially true (and I leave this question open for the time being), it would mean that the study of writing is not a peripheral but a central linguistic task. As a slowly establishing discipline, *grapholinguistics* is the direct response to these considerations.

*Grapholinguistics* is the name of "the linguistic sub discipline [sic] dealing with the scientific study of all aspects of written language" (Neef 2015: 711). The term is a translation of German *Schriftlinguistik*, which was first employed by Nerius & Augst (1988) and adopted by Dürscheid (2016) as the title of her seminal textbook.[2] I follow Neef, Sahel & Weingarten (2012ff.) as well as Neef (2015) in using this term instead of any of the numerous available alternatives, such as *grammatology* (Gelb 1969; Daniels 1990, 2009; in a different sense Derrida 1967), *graphonomy* (Hockett 1951; Daniels 2018), *writing systems research* (the title of a past prominent journal in the field, see below), or *graphem(at)ics*. *Grapholinguistics* not only aligns terminologically with designations used for other linguistic subdisciplines such as psycholinguistics and sociolinguistics but also originates and thus reflects the German-language research area's decades-old practice of acknowledging and investigating writing and written language in their own right rather than as secondary and peripheral matters of linguistics. Additionally, unlike the alternative designations listed above, it is not already occupied as a label for other disciplines or endeavors.[3]

What exactly is grapholinguistics? What does it cover, who are scholars engaging in it? According to Neef, grapholinguistics is a linguistic field dealing with all writing-related issues. This is echoed in an influential German linguistic dictionary, where *Schriftlinguistik* is defined as a "synoptic label for efforts of gaining consistent descriptions and analyses of the written form of language and of developing them to a general theory of writing as a constitutive part of a general theory of language" (Glück 2016d: 596, my translation).[4] While at this point, descriptions of writing abound, the theory of writing mentioned in this definition

2 A French term that was suggested is *scripturologie* (cf. Klinkenberg & Polis 2018).

3 *Grammatology*, although it was first used in a grapholinguistic sense by Gelb (1952), is most strongly associated with Derrida's (1967) philosophical theory of the same name, while *graphonomy* (or also *graphonomics*) is used as the designation of a field that deals with handwriting analysis. *Graphem(at)ics* is problematic since it, at least in the model of writing systems proposed here, only refers to a subdiscipline of grapholinguistics – the one devoted to the graphematic module (cf. Chapter 2). In mainstream linguistics, however, it is commonly interpreted as a *pars pro toto* designation equivalent to grapholinguistics (cf. e.g. the Wikipedia page for graphemics, `https://en.wikipedia.org/wiki/Graphemics`, March 12th, 2020). *Writing systems research* is the only name that is not already occupied and that would suit the discipline. However, although its focus on writing systems is obviously justified, the term insinuates a narrower scope than what is actually studied by grapholinguistics: for example, solely graphetic research endeavors, such as studies that test which connotations or emotions different typefaces evoke, are definitely grapholinguistic but not about the writing system *per se*. Such questions might not always be seen as *writing systems* research.

4 „Zusammenfassende Bez. für Bemühungen, konsistente Beschreibungen und Analysen der geschriebenen Sprachform von Spr. zu gewinnen und sie zu einer allgemeinen Schrifttheorie als konstitutivem Bestandteil einer allgemeinen Sprachtheorie zu entwickeln."

is still lacking. As Dieter Nerius, who is considered the term's founder, noted, it is intended to highlight the integration of the treatment of writing into linguistics (cf. Dürscheid 2016: 12). Against this background, however, *Schriftlinguistik*, while somewhat established, is still notoriously underrepresented both as a term and, more generally, as a linguistic subdiscipline. Yet, as Dürscheid (2016: 11) observes, the situation is improving: for example, a grapholinguistic dictionary (the above-mentioned Neef, Sahel & Weingarten 2012ff.) is in the works as part of a series of dictionaries that also features dictionaries for undoubtedly established linguistic subfields such as morphology and syntax. In general, grapholinguistics now more often appears alongside other linguistic disciplines; for example, it increasingly occurs in linguistic textbooks (e.g. the chapter by Berg & Evertz 2018). Workshops and conferences are devoted to it and, finally, it is treated by an increasing number of courses in linguistics programs around the world.[5] However, while these steps appear like a promising leap in the right direction, they are not a reason to be overly hopeful, as grapholinguistics is nowhere close to representing a coherent field. Interestingly, in this respect, as already implied above, the English-language community lags far behind the German-language community (see below). Since there is little perception of a coherent field, drafting a historical reconstruction is understandably challenging.[6] Nonetheless, I will attempt to sketch some of the relevant milestones in the slow but steady development of grapholinguistics and a scientific community (or better: scientific communities) devoted to writing (cf. also Haralambous 2020).

It is unsurprising that *Schriftlinguistik* is a more widespread term than the English equivalent *grapholinguistics*,[7] a terminological observation that only

5 A quick (and of course non-exhaustive) search for grapholinguistic classes in the German-language area revealed that introductory courses to *Schriftlinguistik* (and sometimes, additional courses focused on specific phenomena of writing) are or were at one point in the recent past offered in, among others, the universities of the following cities/areas: for Austria – Graz, Innsbruck, Vienna; for Germany – Aachen, Berlin, Braunschweig, Chemnitz, Duisburg-Essen, Hamburg, Hildesheim, Jena, Kassel, Köln, Leipzig, Mainz, Oldenburg, Osnabrück, Saarland, Wuppertal; and for Switzerland – Basel, Zurich. Internationally, I found introductory classes in Athens, Austin, Boston, Toronto.

6 While a multitude of histories of writing exists, there are virtually no histories of the study of writing, with only few exceptions: firstly, a handbook chapter by Schlieben-Lange (1996), which, however, covers the 20th century only sparsely and was published before the 21st century that proves so central for grapholinguistics, and secondly, an extensive overview in Spitzmüller (2013: Chapter 3) that interested readers can consult. Both are written in German. A more recent brief overview is given in Haralambous (2020). Historical grapholinguistic context and critical readings thereof are also provided in the many reviews of writing-related works by W. C. Watt (e.g. Watt 1989, 1998, 2002).

7 The English term *grapholinguistics*, in fact, has seldom been used before Neef (2015), and this is only now changing given that it was adopted for the titles of the conference series *Grapholinguistics in the 21st Century* and the book series *Grapholinguistics and Its Applications*. As an adjective, it is found in Seidenberg's (2011) '*grapholinguistic* equilibrium hypothesis'. One earlier use I found (albeit hyphenated as *grapho-linguistics*)

reflects that the field's prominence varies in different linguistic realms. In the German-language area, especially in Germany, linguists have treated aspects of writing and written language as valuable research objects since roughly the second half of the 20th century. At this point, one might argue that some of the most important – and earliest – systematic monographs on writing, including Gelb's *A Study of Writing* (first published in 1952, cited here in its second edition/third impression, Gelb 1969) and Diringer's *The Alphabet* (first published 1948), among others, were written in English. And indeed, to this day, the most relevant books on writing and writing systems are published in English (Coulmas 1989, 2003; Sproat 2000; Rogers 2005; Sampson 2015; Daniels 2018). The same is the case for the central edited volume (Daniels & Bright 1996), an edited four-part collection of some of the most important articles and chapters on writing (Moseley 2014), an encyclopedia (Coulmas 1996a), and a comprehensive bibliography (Ehlich, Coulmas & Graefen 1996). Considering monographs alone, there is nowhere near this breadth of grapholinguistic overview works in German. Why, then, can it be argued that the German-language community has been more instrumental in the development of grapholinguistics? This argument hinges predominantly on the concepts of scientific 'community' and institutionalization. Also, most grapholinguistic research published in English differs in its epistemological interest from German research.

As mentioned, in the German-language realm, acceptance of writing as a linguistic subject and, probably of greater importance, a growing interest in actually investigating it in a systematic fashion, were strong: In her sketch of the development of grapholinguistics, Dürscheid (2016: 12) mentions two influential German research groups devoted to writing. They originated in the different parts of then-separated Germany: the *Forschungsgruppe Orthographie* (founded in 1974 in the GDR, cf. Nerius 2012), and the *Studiengruppe Geschriebene Sprache* (founded in 1981 in the FRG, cf. Günther 1993a; Giese 1993). Some of the most influential "grapholinguists" were members of these groups, including, for the former group, Dieter Nerius and Gerhard Augst, and for the latter, Florian Coulmas,[8] Konrad

is by Penny Platt: "Graphic images are part of a visual vocabulary which has intense personal meaning to the child. There is a symbolic relationship among drawing, writing, reading, speaking, and listening. The meshing of all these processes rightfully belongs to a new science which I have named grapho-linguistics" (Platt 1977: 263). This, obviously, is not the reading of *grapholinguistics* that is advocated here. Another little-received use is by Anthony W. Sariti in a paper (Sariti 1967) covering the Chinese writing system, entitled *Chinese grapholinguistics* and based on his master's thesis. He notes that the "Grapholinguistic System (GS) is concerned with what we usually call the 'written language', that is, it is manifested not by acoustic but by graphic events (writing)" (Sariti 1967: 3, emphasis in original). From this quote, it seems that he uses *grapholinguistic system* in a way that is synonymous to *writing system* as used in this book.

8 Coulmas is one of the few members of this group who frequently also published and still publishes in English although he is originally a member of the German-language grapholinguistic community. This, to some degree, opened up topics treated in the German community to an English-language audience. Additionally, Coulmas also published in French, cf. his chapter in Catach (1988). His main interests are writ-

Ehlich, Hartmut Günther, Peter Eisenberg, and Otto Ludwig. These scholars are all linguists or Germanists and, crucially, writing and written language are not peripheral subjects in their respective lists of research interests – they are central. As Günther (1990a) notes, in the years leading up to 1990, more than 15 habilitation dissertations submitted in Germany were devoted to aspects of writing and written language.[9] Based on this, he concludes that a paradigm shift had occurred, with linguists feeling they no longer needed to justify their interest in writing (echoing Baroni's programmatic statement quoted above). This development initially culminated in the two volumes of the handbook *Schrift und Schriftlichkeit* (*Writing and its Use*), edited by Günther and Ludwig with the support of the remaining members of the group *Geschriebene Sprache* and published by De Gruyter in 1994 and 1996, respectively (cf. Günther & Ludwig 1994, 1996). The handbook boasts over 140 chapters written in German and English and deals with a vast spectrum of aspects of writing. It unites a great number of scholars interested in the field, including even former "rivals" from the *Forschungsgruppe Orthographie* such as Dieter Nerius as well as non-German scholars. Strikingly, however, the term *Schriftlinguistik* does not occur even once throughout the entire handbook, implying that even though for the first time, an extremely elaborate handbook was devoted to the subject of writing, there was still no perception of a coherent field studying it – yet.

The second major difference between the Anglo-American and German treatments of writing mentioned above is that in the latter, the topic is approached in a more theoretical, explanatory matter, whereas in publications by English-speaking scholars, the focus is usually primarily on description. This coincides with a focus on synchrony in the German grapholinguistic tradition and diachrony in the Anglo-American tradition. Works such as Rogers (2005), Sampson (2015), and Daniels (2018), and also the contributions in Daniels & Bright (1996), offer excellent descriptions of a great variety of the world's writing systems, but when they broach theoretical linguistic aspects of writing as a general phenomenon, they fall short of reaching the level of sophistication characteristic of the German grapholinguistic tradition. Notably, however, this was also not their main goal. Most strikingly, these works do not offer methods or categories for further productive linguistic analyses of writing systems (a theoretical chapter in Rogers 2005, a book that is advertised as a textbook, is an exception here). By contrast, German grapholinguistics has focused heavily on these theoretical aspects, asking how writing systems *can be* described instead of 'just' describing them. However,

ing systems other than German, including Japanese and Chinese. Notably, his 2003 monograph (published in the *Textbooks in Linguistics* series by Cambridge University Press), in its epistemological interest, its makeup, and its aims, resembles the other English-language descriptive treatments of writing systems and refers to little of the theoretical footing previously established in German grapholinguistics (and partially by himself, cf. Coulmas 1989: Chapter 3).

9 This tradition continued, as between the early 2000s and today, German linguists Martin Neef, Nanna Fuhrhop, Ursula Bredel, and Kristian Berg, among others, also devoted their habilitation dissertations to aspects of writing and written language.

ironically, in the process, it has sacrificed a greater, more universalist horizon, restricting itself to the analysis of alphabetic writing systems, predominantly German, on which a large portion of its theorizing is based (cf. Fuhrhop 2018). Ironically, a mixture of the two approaches appears most fruitful: careful descriptions of vastly diverse systems can only inform theoretical work, including models, concepts, and, in general, abstractions. Sadly, there is not only a massive (likely language-induced) lack of reception of German literature on behalf of English-speaking scholars[10] but, less understandably, readily available English literature is also often not considered in German works. Not to mention the lack of reception of grapholinguistic findings from other communities, such as the French, where, in 1962, Nina Catach founded the research group *Histoire et structure des orthographes et systèmes d'écritures*, which very productively investigated writing and written language. In sum, one could conclude that there exists an abundance of valuable grapholinguistic research that is scattered around various academic cultures, published in different languages, and across various academic fields. This research is waiting to be integrated into a coherent overall picture – which is a little bit like putting together a large puzzle (and can impossibly be achieved by a single person).

Even though *Schriftlinguistik* was not a label the above-mentioned handbook *Schrift und Schriftlichkeit* (Günther & Ludwig 1994, 1996) put on itself, it already foreshadowed what Dürscheid would later subsume under the term *Schriftlinguistik* in her textbook *Einführung in die Schriftlinguistik* (first published in 2002, cited here in its fifth edition published in 2016), which proved constitutive for the discipline. In fact, the handbook's preface to this day offers what I would argue is the best characterization of grapholinguistics found in the literature:

> Due to the diversity and heterogeneity of the subject areas [covered by the handbook, D.M.], a number of different scientific disciplines need to be involved in studying them: philosophy and anthropology, linguistics and literary studies, sociology, psychology, education, history – to mention merely a few. [...] The form which research takes in a given discipline reflects the theories and methods relevant to the respective field; the findings are thus tied to these individual theories and methods. Each discipline studies a given aspect of *Writing and Its Use*, and a relatively complete picture can only emerge when all of them are combined in some way. In this sense, *Writing and Its Use* is an interdisciplinary subject, and research needs to take this into account. [...] the study of *Writing and Its Use* has been restricted to isolated research interests of the individual scientific disciplines. *Writing and Its Use* has thus never become a research subject in its own right, which is why as yet there is neither a unified theory nor has there been an interdisciplinary exchange of theories, problems, and research methods.
> (Günther & Ludwig 1994: XXVIII f., emphasis in original)

**10** There are exceptions to this: some works of the German-language community were dispersed in English, for example in the edited volumes *Writing in Focus* (Coulmas & Ehlich 1983) and *New Trends in Graphemics and Orthography* (Augst 1986), to which not only German grapholinguists contributed but also scholars from other academic cultures and regions working on various aspects of writing, among them Jack Goody, David R. Olson, Margaret Martlew, William Haas, Max Coltheart, and Nina Catach.

If, in this quote, *Writing and Its Use* is substituted for *grapholinguistics*, this assessment perfectly sums up not only the state grapholinguistics was in two decades ago but also its current state.[11] The many disciplines that Günther & Ludwig list are partially echoed in the subjects Dürscheid treats in her textbook, which features chapters on the relationship between spoken and written language, writing system typology, the history of writing, graphematics, orthography, typography, and reading and writing acquisition. While this does not approach the breadth of topics covered by the above-mentioned handbook given that the textbook's focus is clearly linguistic, it still reflects the interdisciplinary nature of the field. Why, then, if we return to the question of terminology, is grapho*linguistics* a fitting designation instead of one of the more neutral alternatives graphonomy, grammatology, or writing systems research? As I will argue below, writing, following a narrow definition, refers only to those graphic (i.e. visual and/or tactile) "marks" that represent language. This excludes marks that refer (directly) to ideas or extralinguistic referents. Writing is always intimately tied to language, and language is the subject of linguistics. The term *grapholinguistics* highlights this linguistic basis while the compositional analogy to terms like psycholinguistics and sociolinguistics underlines the inherent interdisciplinarity and implies that grapholinguistics does not exclude findings, methods, and theories from other disciplines.

I want to propose an additional distinction for grapholinguistics that is common in other linguistic subdisciplines: *theoretical* vs. *applied grapholinguistics*. Whether this division is necessary or useful is debatable, and in fact, this book is located somewhere between these two poles, if palpably more on the theoretical end of the spectrum. Theoretical grapholinguistics deals with the theory of writing, which includes the investigation of the relationship between speech and writing (and, although to this day to a significantly lesser degree, sign language and writing), the construction of models of writing, and the development of descriptive tools for the unified description of writing systems. Its immediate subbranches are *graphetics*, *graphematics*, and *orthography research*.[12] Applied grapholinguistics, on the other hand, deals with but is not reduced to questions of how this theoretical knowledge can be put to use for "writing-related real-life problems".[13] This includes educational, psychological, and medical questions such as literacy instruc-

11 This is echoed in Judson's (2017) review of a more recent volume dealing with historical graphematics and its 'concepts and methods' (Cotticelli-Kurras & Rizza 2017), as she implicitly criticizes the lack of comparative and interdisciplinary grapholinguistic theory: "What the introduction to this volume lacks [...] is a broader theoretical or methodological discussion of the book's potential interdisciplinary impact: how developing methodological approaches to or studying particular aspects of one writing system may help to illuminate others."

12 In German, *Orthographie* designates not only the research object 'orthography' (in the sense of a prescriptive regulation of writing systems, cf. Chapter 3) but also the field studying it. In English, using *orthography* for a field of study seemed odd to me, so I felt compelled to add *research*.

13 This is a slight modification of Wikipedia's definition for *Applied linguistics* (cf. https://en.wikipedia.org/wiki/Applied_linguistics, March 14th, 2020).

tion or the diagnosis and treatment of disorders that affect reading and writing. The theoretical findings of this book could be relevant for applied grapholinguistics insofar as they will hopefully be of use not only for grapholinguists but for people working on any of those real-life writing problems as well.

To conclude this brief historical sketch of grapholinguistics, I want to arrive at the present and, even more importantly, take a look at the future by asking: what is the current state of grapholinguistics? It might be pessimistic to phrase it this way, but the discipline's (first) 'heyday', if it ever really experienced one, appears to be over: to my knowledge, there is (save for some specific funded projects on writing) no contemporary equivalent to the writing-related research groups named above, and not all of the members of these groups are still active or as active as they previously were. And although steps are taken in the 'right' direction, the question is whether they suffice. As of yet, there exist virtually no chairs for grapholinguistics[14] and there are no linguistics programs devoted to it.[15] This lack of institutionalization results in the fact that many of the findings that would be very important for the discipline detrimentally continue to be scattered across the many fields involved.

Even today, I believe it is hard to find people who self-identify primarily as scholars of writing – technically, if you will, 'grapholinguists'. Scholars such as Peter T. Daniels[16] or the late Earl M. Herrick might be exceptions to that. In linguistics, expertise in the field of writing systems is commonly perceived as icing on the cake, not as a serious and sufficient research interest in itself. I experience this regularly when I am asked *what other* linguistic interests I have besides grapholinguistics. In response to this, I argue: grapholinguistics is not a hobby. It is not a sideline. It is a serious field of study and a subdiscipline of linguistics. This is highlighted by the fact that to seriously engage in it, broad knowledge of a variety of other subfields of linguistics is necessary, reflecting that writing interacts with all levels of language. The present book is a direct product of this thinking.

Ultimately, the picture is not quite as grim as the preceding paragraphs paint it: first of all, with a high-ranked journal solely devoted to writing,[17]

---

**14** In fact, I know of only one grapholinguistic chair, which was advertised as of March 2020 at the University of Hamburg.

**15** As of June 2016, there was a plan for a Master's program at the University of Haifa in Israel, entitled *Literacy Development and Writing Systems*. It is/was planned as an "inter-disciplinary approach to the study of literacy learning across languages and writing systems" that bridges psychology and linguistics (David Share, p. c.).

**16** In an *About the author* section following one of Daniels' articles, he is described as "one of the few linguists in the world specializing in the study of writing systems" (Daniels 2006b: 45), which underlines how rare of a species grapholinguists are.

**17** In earlier drafts of this book, an additional journal was listed here: For ten years, starting in 2009, *Writing Systems Research* was devoted to grapholinguistic topics. Unfortunately, in 2020 it was announced that the journal was discontinued after volume 11 (2019) (cf. https://www.tandfonline.com/toc/pwsr20/current, July 24th, 2020).

*Written Language and Literacy*[18] (founded in 1998), and additionally the (yet) too little-known *Scripta*[19] (also since 2009), not to mention several writing-related journals from other disciplines involved in grapholinguistics,[20] there now exist official and highly visible ways of disseminating the disciplinarily heterogeneous findings that can be subsumed under the heading of grapholinguistics. Furthermore, at least five series of conferences focus on writing: the international workshops of the *Association of Written Language and Literacy*, a series of workshops entitled *The Idea of Writing*, proceedings of which have been published in an eponymous series by Brill, a conference series entitled */gʁafematik/* whose proceedings are published in a new open-access series entitled *Grapholinguistics and Its Applications* (which this book is a part of), another one named *Laut Schrift Sprache* (English title: *Script and Sound*), and the international conferences of the *Hunmin jeongeum Society*, which is also the publisher of *Scripta*. Additionally, an increasing number of workshops or sessions at general linguistic conferences are devoted to the topic.[21] This creates a welcome opportunity for scientific exchange. To this day, Germany remains the heart of the international grapholinguistic community, with scholars in Cologne (Martin Evertz, Frank Kirchhoff), Oldenburg (Nanna Fuhrhop, Franziska Buchmann), Bonn (Kristian Berg), Osnabrück (Karsten Schmidt), Düsseldorf (Stefan Hartmann), Halle (Florian Busch), Hildesheim (Ursula Bredel), Braunschweig (Martin Neef, Miriam Balestra), Hamburg (Jannis Androutsopoulos), Bochum (Sven Osterkamp, Gordian Schreiber), Bamberg (Renata Szczepaniak) working on questions of writing and written language, to name only a few (and most certainly leave out many). Important publications which are expected to become central to the field keep appearing, such as Daniels (2018), a culmination of his scholarship from almost three decades,[22] or Domahs & Primus (2016),

18 https://benjamins.com/catalog/wll (March 14th, 2020).

19 http://www.scripta.kr/ (March 14th, 2020).

20 Examples are *Reading and Writing* and *Scientific Studies of Reading*. These journals are not exclusively or even primarily linguistic, but given their focus on the production and perception processes of writing and reading, what is published in them is pertinent grapholinguistic research. This is reflected by these journals' self-proclaimed commitment to interdisciplinarity, which is at the heart of grapholinguistic research. Another example is *Visible Language*, an interdisciplinary design journal focusing on visual communication research. It has published a breadth of grapholinguistic literature, among which are some of Earl M. Herrick's central contributions to the field (e.g. Herrick 1974).

21 In this vein, workshops such as *Written and spoken language as modalities of one language system* (Conference of the German Linguistic Society, February 2016), *Theories and methods of grapholinguistics* (Austrian Linguistics Conference, November 2016), or *The evolution of writing systems: Empirical and cross-linguistic approaches* (Conference of the German Linguistic Society, Hamburg 2020) are worth mentioning.

22 Note that Daniels' work remains descriptive and historical. It informs grapholinguistics greatly, but it is not *per se* grapholinguistic. Daniels (1991) foreshadowed this himself with his rejection of a 'structural graphemics' and with it the idea of investigating writing by using the same methods and concepts that are employed in the study of language. Against this background, calling his work 'grapholinguistic' would con-

a handbook that places written language alongside spoken language and sign language and weights all of them equally. Finally, as becomes evident when attending relevant conferences, there is a young generation of scholars interested in writing and numerous grapholinguistic doctoral dissertations are in the works that will hopefully also further flesh out the field.

Speaking of doctoral dissertations, this book is a modified version of one. The motivation that underlies it originates from the above-mentioned observation that the state of the field is fragmentary and unorganized. Consequently, at its core, it is an attempt to reconcile different grapholinguistic cultures to start building a grapholinguistic theoretical framework. Given this main goal, it is not innovative in that it presents new revelatory data about writing but in that it adopts broader integrative perspectives on the subject of writing and asks questions of universal nature to arrive at explanation rather than just settling for description. Indeed, many facets of writing have already been illuminated brilliantly, including a large portion of its history and development or the structural description of many major (and some minor) ancient and modern writing systems of the world. However, as implied above, the promising results of individual achievements remain mostly unconnected to related findings, which prevents an exploration of the bigger picture and is detrimental to the development of a theory of writing. This starts with a lack of basic concepts and terminology that would make possible a unified description and, crucially, a comparison of different writing systems. Notably, the concepts and models that do exist are largely Eurocentric (cf. Yan 2002), admittedly an old shoe of a criticism that, alas, still fits, as Share (2014) observes for contemporary psychology, in which he speaks of an "alphabetism in reading science". This aftertaste of an alleged superiority of the alphabet remains tastable not only in psychology but also in other grapholinguistic areas.

The lack of a grapholinguistic theory has been lamented before. Roughly two decades ago, in a review of the – to this day – arguably most important edited volume on the world's writing systems (Daniels & Bright 1996), W. C. Watt, a fervent observer of writing-related research, expressed his wish for 'more' and explained what that entailed:

> 'More' would constitute, or at least contribute to, a semiotic theory of writing systems: a theory that would explain, to put it pithily, why each such writing system is the way it is, instead of some other way, and why all such systems have in common what they have in common. Such a theory might begin by examining the nature of the problem facing anyone who wishes to record a spoken language, competently addressing the question of how speech is most naturally segmented in relation to how it is most naturally or most easily translated into visible marks. Such a theory might proceed to examine the purely physical factors that deter-

tradict his assessment. Ironically, his research is ideally combined with an unlikely complement: the very structural 'graphem(at)ics' that he rejected and that is encapsulated, for example, in Fuhrhop & Peters (2013). This structural graphematics, ironically, suffers from a narrow typological (read: alphabetocentric) horizon that Daniels' comprehensive and minute descriptions of a broad range of writing systems could help widen.

> mine the forms of writing systems: the nature of available writing materials and the nature of the pertinent human physiology, especially that of hand and eye. […] Such a theory might continue by examining the cognitive factors that determine the forms of writing systems: the human disposition to generalize and/or to misremember in certain directions rather than others, for instance. Any such theory would, of course, if successful at the tasks just set forth, also largely explain how one writing system can 'evolve' into another […]. (Watt 1998: 118)

Watt's observations call for a shift from description to explanation. In other words, it does not suffice to describe certain aspects of writing or even entire writing systems without then integrating the findings into a larger theoretical framework to gain explanations of how the described structures came to be. It is this very framework that is still largely missing and that the present book is devoted to.

Given the abundance of writing-related research, the first step of constructing a theory of writing has already been taken. The next necessary step is to broaden the horizon. Writing systems of vastly different languages, both genetically and typologically, must be considered, and it must be evaluated if and how comparative categories can be established that suit all of them. The comparison that such categories facilitate will reveal not only differences but also commonalities between diverse writing systems. Concerning the latter, universal tendencies are of greatest relevance. To explain why they exist, one must turn from the mere structure to the use of writing, i.e. the acts of writing and reading, and analyze them both from a psychological and a sociocommunicative perspective. The structure of scripts as inventories of visual marks (e.g. Cyrillic script) and writing systems, the systems relating these visual marks to a given language (e.g. the Russian writing system, which relates Cyrillic script and the Russian language), is subject to human pressure. This means that it is heavily influenced by the fact that prototypical readers and writers are equipped with brains, eyes, hands, and a need to communicate. Accordingly, in this book, it is argued that the structure of writing is a reflection of how humans have shaped it ever since its inception some thousands of years ago. In the establishment of a theory of writing, it is thus paramount to reconstruct how human conditions and writing systems interact exactly. Here, given that language systems and writing systems are both semiotic phenomena and exhibit myriad parallels, a theory of writing need not be developed from scratch but can build extensively on existing theories of language, especially functionalist and usage-based theories which rely on extralinguistic factors and external evidence. One such theory in particular will lend the prospective grapholinguistic theory many of its ideas and concepts: Naturalness Theory.

Naturalness Theory is a theoretical framework consisting of several approaches that share the core view that the nature of humans – subsuming, among other things, their physiology, cognition, and socialization – shapes the structure of language. Natural Phonology, the branch that introduced the theory, claims the existence of processes that facilitate articulation and perception of speech by eliminating difficulties. These processes are considered and termed *natural* as they are based on human physiology. In other words, what is easy to speak with our mouths, lips, etc., to hear with our ears, and to process with our brains is interpreted as natural. A second influential branch of the theory, Natural Morphology, transferred these core ideas to morphology and proposed that the semiotic structure of morphemes – as dyadic signs – bears on human cognition. This led to a list of so-called naturalness parameters that describe aspects of the semiotic structure of morphemes claimed to be relevant in cognitive processing. Ultimately, in both Natural Phonology and Natural Morphology, *natural* pertains to human nature and is a scalar, evaluative attribute that is ascribed to those features of language that are relatively easier to process than others. Note how Watt, in his quote above, likely oblivious to Naturalness Theory, uses the words "naturally" and "nature" in ways that fit perfectly into a naturalist paradigm. Naturalness Theory shares commonalities with many other linguistic theories, including Markedness Theory, Optimality Theory, and usage-based approaches, but in its striving for an extralinguistic explanation for linguistic data, it serves as an especially suitable basis for a grapholinguistic theoretical framework. An extension of Naturalness Theory to the subject of writing is predicted to be capable of achieving many of the points Watt lists as desiderata – in a nutshell, to explain "why each writing system is the way it is". However, I want to note upfront that I am not a fervent proponent of any theory, including Naturalness Theory. Primarily, I am a linguist interested in writing, and for reasons explained throughout this book, Naturalness Theory appears to be a suitable framework. However, when shortcomings of the theory are encountered, they will be addressed.

The fact that Naturalness Theory is a promising framework for the study of writing has formerly been pointed out by Munske (1994) and Baroni (2011). Their interesting preliminary ideas, however, have not yet been developed into a full theory, which will change with this book. In the vein of Stampe's (1979) constitutive work on Natural Phonology, in which he outlined roughly the central tenets of the approach, and Dressler's (1989) sketch of the core semiotic parameters relevant in Natural Textlinguistics (yet another, if minor, branch of the theory), this book will provide the most important cornerstones in the first sketch of a functionalist grapholinguistic theory – *not* a description of a full-fledged theory. To achieve this, as was done by Stampe (1979) for phonology and Dressler (1989) for textlinguistics, examples from a variety of different systems will be given and, in the vein of naturalist methodology, extralinguistic evidence from different sources – including literacy acquisition and the diachronic development of writing – will be considered. The main focus is on explaining universal tendencies; however, typological aspects specific to certain writing systems will also be discussed. What also needs to be disclosed is that most examples stem from modern writing systems, which potentially implies that the theory was synchronically oriented. This is true only partially: the theory aims to explain why all writing systems are the way they are, and this includes ancient writing systems that are no longer in use. The importance of diachrony is underlined by the fact that one of the most important types of explanatory evidence is, as mentioned, the historical development of writing systems. The focus on examples from modern writing systems results from the simple fact that for them, other types of evidence – such as highly relevant psycholinguistic evidence – are also available.

The scope and aims of this book might appear enormous and (overly) ambitious, making it reasonable to explicitly formulate its main questions, the central of which is: What are the cornerstones of an explanatory rather than a 'merely' descriptive grapholinguistic theory, i.e. what allows the shift from the question of how a writing system is structured to why it is structured that way? Which are the central concepts of such a theory? Which methods must be adopted and which data consulted to construct it? The investigation of these questions can result in a tentative answer to the question of which forms explanation in grapholinguistics might assume. If successful, this endeavor will also indicate roughly how to proceed to gain further insight into the nature of writing. But, to put it in terms of writing: when it comes to a theory of writing, this book does not provide the fine lines, it only offers the first rough broad strokes.

A few more limitations must be addressed in advance. Firstly, the presented sketch of a theoretical framework will be preliminary as it is impossible to take into account all of the world's writing systems in one study. However, based on examples from a variety of different systems, a rough framework for the comparison of writing systems in the form of a *tertium comparationis* will still be made available for future research. In this context, Rogers' (1995: 31) provocative claim comes into play, namely that "some writing systems are better than others", which, as he added, "could certainly be debated". This debate, however, at least one objective in nature and untainted by the presumption of the alphabet's superiority, has largely failed to materialize. This is likely precisely due to the lack of a tertium comparationis. Crucially, what this book will *not* be is a minute description of any specific writing system, nor a detailed comparison of any two writing systems. Instead, it will be an investigation of the categories relevant in a theory of writing, and these represent relevant tools not only for seeking explanations in single writing systems (e.g. "Why is the German writing system structured the way it is?") but also for the comparison of writing systems with respect to certain relevant features (e.g. "Is the grapheme inventory more transparent in the writing system of Thai or Chinese?").

Secondly, this book is inevitably shaped by my background. I am a generalist asking broad theoretical questions and aiming for a bigger picture rather than specializing in any given writing system. Although I criticized these biases above, my research might still implicitly be Eurocentric or even Germanocentric. It is central to be aware of these restrictions. A theory of a highly complex phenomenon such as writing cannot be constructed by a single person from a single field. Where one's expertise ends, other scholars – experts on specific writing systems, scholars from other fields such as psychology, the cognitive sciences, history, etc. – must step in to clarify or fill in the blanks, lucidly highlighting the necessity of interdisciplinarity in grapholinguistics. For this reason, in the course of the book, pressing desiderata will be identified as it will become obvious where research is still lacking. Ultimately, a sketch of a theory of grapholinguistics is simultaneously a diagnosis of the current state of the entire interdisciplinary field as it identifies those underrepresented areas in which both further research and increased exchange between different disciplines are paramount.

The core of this book is divided into two parts which are dedicated to description and explanation, respectively. Accordingly, Part II is a detailed treatment of descriptive structural grapholinguistics. Its aim is to establish a structural and methodological framework that can be applied to all – as opposed to only alphabetic – writing systems. Its subdivision into three chapters is based on Neef's (2015) multimodular model of writing systems, which is presented at the outset of Part II and, in essence, holds that writing systems are based on an underlying language system and consist of a script, a graphematic module, and – optionally – an orthographic regulation. Thus, Chapter 1 is devoted to the graphetic module as well as graphetics, the grapholinguistic subdiscipline studying it. In this vein, various aspects of the description of the materiality of writing will be discussed. Following Meletis (2015, accepted a), different subdisciplines, levels of description, and relevant graphetic concepts and units will be presented. Additionally, the question of what a typology of scripts – i.e. inventories of visual shapes such as Roman script, Cyrillic script – could look like will be addressed. Next, Chapter 2, the most extensive portion of this book's descriptive part, delves into the graphematic module – the one that relates the visual to the linguistic. One of the central issues of graphematics concerns the relationship between language, speech, and writing. Accordingly, open questions pertaining to this problem and notions of dependence vs. autonomy will be discussed. A problem closely related to this question that will also be treated here is how the infamous concept of *grapheme* can be defined in a broad manner to make it applicable to all types of writing systems. Furthermore, valuable contributions of German grapholinguistics in the investigation of other graphematic units – such as the graphematic syllable, word, or sentence – which were previously only available in German will be made accessible to an international audience and simultaneously critically scrutinized with respect to their cross-linguistic applicability. Finally, as is it based mostly on the canonical graphematic relations in different writing systems, the well-established field of writing system typology will be treated by discussing some of the core typologies that have been proposed in the past. As the final descriptive chapter, Chapter 3 is dedicated to the orthographic module. In the English-language realm, partially based on the fact that English itself has a self-organizing orthography, *orthography* is often treated as a synonym of the descriptive term *writing system*, which obscures the normative and sociolinguistic character of orthographies as central concerns of linguistic policy. The chapter thus describes the properties of orthographies as standardizations of different types of writing systems as well as their relevance as sociolinguistic phenomena.

In Part III, the focus is shifted to explanation. Its introduction deals with several core ideas from Naturalness Theory and details how they can be fruitfully repurposed in a theory of writing. Notably, these ideas include a hierarchical organization of the theory in system-independent (i.e. universal), typological, and system-dependent parts and the consideration of external, i.e. extralinguistic, evidence. Furthermore, to establish the perspectives that will prove crucial for a grapholinguistic theory, previous attempts at evaluating writing systems, especially collections of relevant criteria, will be discussed. Those criteria, which recur frequently in scholars' lists, will be condensed and assigned to four categories that are evaluatively interpreted as 'fits', i.e. ways of how writing systems meet the different demands imposed on them, and are subsequently treated in dedicated chapters. Internally, these chapters are subdivided according to the three modules of writing systems presented in Part II (graphetics, graphematics, orthography), investigating how these modules satisfy the requirements of a given fit. The *systematic fit* (Chapter 5), as the name suggests, evaluates how good of a system the given modules are, which is primarily done by assessing how systematic the relationships are between their units as well as how many features are required to describe said units and how evenly these features are spread throughout the system. Crucially, the systematic fit is the only fit that can be analyzed purely system-internally. To illustrate it, the main focus will be on the graphetic module, specifically scripts which, as visual systems, provide writing systems with their substance. The central question here is how the relationship between the individual visual shapes of scripts can be systematic – or not. The chapter will also highlight exemplarily how the systematic fit can be evaluated for the graphematic and the orthographic modules. Next, for the *linguistic fit* (Chapter 6), we turn primarily to the graphematic module to explore how well writing systems suit the language systems they are based on. For example, whether the writing system of German, given its structure and properties, is a good fit for the German language. This is possible on the basis that writing systems are semiotic systems consisting of signs of writing that relate visual units with linguistic units (such as phonemes and morphemes). The linguistic fit, then, assesses the semiotic structure of these signs of writing, i.e. the semiotic relationship between their visual and linguistic constituents concerning several parameters such as transparency and uniformity. In sum, these parameters serve as tools to test the validity of claims such as "every language gets the writing system it deserves" (Frost 2012: 266).

While the systematic fit evaluates the internal structure and systematicity of writing systems and the linguistic fit treats the relationship with their respective languages, the final two fits are concerned with the paramount interaction between writing systems and their users. Accordingly, the *processing fit* (Chapter 7) determines how well writing systems meet the demands imposed by human physiology and cognition. For the graphetic module, this primarily concerns visual complexity, whereas, for the graphematic module, it affects the semiotic structures already described in the context of the linguistic fit and the question of how they are processed cognitively. Last but, as will be argued, definitely not least, the *sociocultural fit* (Chapter 8) judges how writing systems satisfy their users' social, cultural, and generally communicative needs. As a cultural technique, writing is deeply entrenched in culture, and accordingly, users hold numerous expectations as to how it should reflect their culture and identity. Notably, the sociocultural fit is the only one for which the orthographic module will be discussed in detail. This is based on the fact that users of writing systems that are orthographically regulated are dealing with literacy predominantly through a normative lens, in other words: they are confronted mainly with *correct* writing. Thus, their sociocommunicative and cultural interactions are not with the entire writing system but primarily with its orthography.

After these four fits have been characterized in detail, a discussion in Part IV puts the findings of the descriptive and explanatory parts in perspective. It not only gives a critical summary but emphasizes the relevance of an explanatory grapholinguistic theory for the field and indicates how it can be useful also in applied settings such as education and type design. The conclusion (Part V) will then collect key take-aways. As the final part of the book, an outlook (Part VI) will sketch the next steps necessary to continue building a theory of writing.

With this book, I hope to contribute to a theory of writing, allowing linguists and interested researchers from other fields to learn more about the fundamental nature of writing as mediated by its users – in other words, by human nature. This, I believe, is also valuable for linguistics on a broader scale: it would be naïve to assume that in literate communities, writing does not exert an influence on speech, which is still regarded as the primary object of interest in linguistics. Writing affects language systems (even if they are primarily spoken) on many of their levels, ranging from phonology to morphology, syntax, and pragmatics. Gaining a better understanding of how writing works and, in turn, how this is connected to human nature possibly allows an integration into larger theoretical frameworks that are interested in language as a whole.

# II Description

Grapholinguistic research is undeniably at its strongest when it comes to description. This is not surprising given that description has been its main focus, which has led to a myriad of treatments of writing systems that provide invaluable insight into their different features. Take the arguably most comprehensive work covering diverse writing systems, Daniels & Bright's fittingly titled *The World's Writing Systems* (1996). It consists of concise and in and of themselves very informative chapters on a vast array of different writing systems written by respective specialists. What is missing, however, is a strong common thread. In this context, the volume actually reflects, on a smaller scale, the state of the entire field: descriptions of writing systems coexist but rarely reference one another. What is emphasized by this situation is the sheer diversity of the world's writing systems. Their unique features, their differences. These, of course, are indeed of the utmost importance and should not be swept under the rug, but they also ought not to obscure the fact that there is also unity to be found. Indeed, even the (superficially) most diverse writing systems will have something in common. In short, what was rarely – if ever – done in work that can be retrospectively labeled as grapholinguistic was taking a step back from individual descriptions to arrive at a framework that is capable of integrating data from all writing systems.

Much like a unified description of distinct and diverse languages faces challenges, so does a framework that aims to be capable of describing all writing systems. As mentioned above, the main impression among scholars appears to be that writing systems are too diverse to be described with the same set of concepts. However, this cannot be claimed when it has not been attempted yet. This is precisely the goal of this part of the present book: presenting a unified descriptive framework that can account for the diversity of writing systems while still allowing a comparison. Firstly, what will be shown in this context is that in the description of writing systems, an astonishing number of concepts established in other linguistic subfields (examples range from the subdisciplines of phonetics to concepts such as allomorphy and phonotactics) can be fruitfully transferred to grapholinguistics, underlining that both writing systems and languages are semiotic systems and share important structural parallels. However, many other facets of writing systems can only be captured with genuinely grapholinguistic concepts, i.e. ones that are not modeled after other linguistic concepts. Secondly, when it comes to these concepts, one must walk the fine line between defining them in a way that is either too specific to account for the diversity of the world's writing systems or too broad to be of real theoretical value – and walking that line is no easy endeavor. To link to a discussion that initiated in linguistics, the question is whether grapholinguistic description should rely on specific descriptive categories or on more loosely defined comparative concepts (cf. Haspelmath 2010). As will become evident in the following chapters, the present approach leans heavily towards the latter. This is driven by the trivial assumption that all writing systems must have a common core so that users are able to use them, and comparative concepts in grapholinguistic description should reflect that. Crucially, this does not mean that details should be discarded, ignoring the (often fascinating) idiosyncratic features of distinct writing systems that are frequently at the center of individual descriptions. The question is, however, which of these details are truly relevant in a comparison of writing systems – which of course depends on the epistemological interest motivating it in the first place. In any case, at this point, it suffices to say that to build a preliminary comparative framework, abstraction and generality are good places to start.

In this vein, a general model of the structure of writing systems is an absolute necessity for a grapholinguistic theory (whether it be descriptive, explanatory, or – optimally – both). For only if we have understood how writing systems are built can we adequately formulate and study specific grapholinguistic questions. This view, however, has not always prevailed, and in fact, most of the research on writing systems does not explicate any respective model of writing (systems) that it is based on. This, one could argue, is because much of this research effectively lacks an underlying model or theory. The starting point for the model that I propose here is Martin Neef's (2005, 2012, 2015) *multimodular theory of writing systems*. This theory aims to describe the subsystems that constitute writing systems, which are named *modules*. A modified version of the model is illustrated in Figure 1: as we see, writing systems are dependent on a language system and consist of the three modules of graphetics, graphematics, and, optionally (see below), orthography. In the following, the model will be presented more thoroughly as its modules and their interrelations are characterized. After that, of course, the bulk of this book's descriptive part dedicates entire chapters to each of them.

A given LANGUAGE SYSTEM such as English is the basis of each writing system, in other words: its core module. According to this model, thus, no writing exists without an underlying language system. Notably, this is not the only possible way of defining writing, but rather an axiom of the so-called *narrow definition of writing* in which, restrictively, only the graphic representation of *language* is interpreted as writing. In this reading, writing is *always* glottography ('language-writing') and *never* semasiography ('meaning-writing', cf. Gelb 1969: 12; Schmitt 1980: 7–11). Indeed, according to the contemporary *opinio communis*, semasiography, the direct graphic representation of thoughts and meaning which is often associated with nebulous and largely abandoned terms such as 'ideography', is generally not acknowledged to be writing.[23] As an example of this module, consider the English language as the core of the English writing system. Like every language, it consists of linguistic units at various levels: phonemes, syllables, etc. as units of sound, morphemes, lexemes, etc. as units of meaning, and phrases, sentences, etc. as larger chunks of language that are constituted by the combination of smaller units. The observation that by default, writing systems predominantly favor one of these linguistic levels supports traditional writing system typology in determining the type of a writing system (cf. Section 2.7). For instance, units of the English writing system primarily represent/correspond with phonemes – consonant as well as vowel phonemes –, making it an alphabet. Notably, this typologization, which is sometimes reductionistically dismissed as 'phonocentrism' as it highlights phonographic relations, does in no way deny that other linguistic levels such as morphology are also of relevance in the writing system.

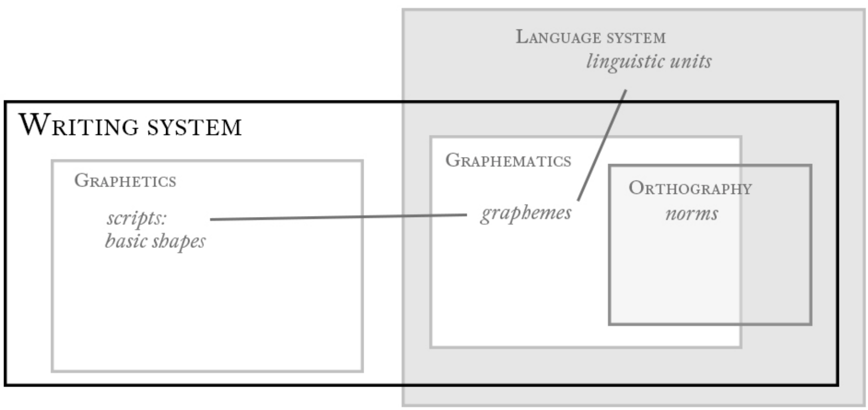



Figure 1. Multimodular model of writing systems, from Meletis (2018: 61)

At the most general level of the typology of writing systems, a categorical distinction is made between writing that represents *sound* (phonography) and writing

23 Unlike ideography, the concept of pictography should not be discarded. It is true that today, in most writing systems, there generally exists no iconic relationship between the visual basic shape that materializes a grapheme and the meaning of the linguistic unit it corresponds with (i.e. a morpheme), as basic shapes – even those which at some point were pictographic – have become increasingly abstract (cf. Section 6.2). This does not mean that the possibility of pictography should be excluded, if pictography is understood as a feature of writing systems and not as its own type of writing systems. Accordingly, a morphographic grapheme – a grapheme that corresponds with a morpheme – can be pictographic when its basic shape visually resembles the concept that the morpheme's signatum refers to, e.g. a tree. What should be discarded, though, is the idea that pictographic graphemes refer directly to the concepts they depict. Essentially, this view succumbs to the same fallacy as the assumption of ideography, i.e. that writing was capable of referring to extralinguistic referents (or 'ideas') directly. There are signs that achieve this – however, they are not considered to be writing but semasiography.

that represents linguistic *meaning* (morphography, logography). Crucially, units of sound such as phonemes and syllables bear no meaning, they only differentiate meaning. Morphemes as well as lexemes, on the other hand, do bear meaning, which means they automatically also differentiate meaning. Simultaneously, these latter meaningful units can be pronounced, they have a phonological representation, i.e. are made up of phonemes. Language, in this sense, is doubly articulated or dually patterned (cf. Martinet 1949, Hockett 1960): comprised of meaningful units that are themselves composed of meaningless discriminative units. This informs grapholinguistic research directly, as will become evident in the discussion of the Chinese writing system, in which the most central graphematic relation is the one between visual units and morphemes – morphemes that, however, due to double articulation, directly correspond with pronounceable syllables. Not seldom did scholars discuss just how "phonetic" the Chinese writing system is, leading up to the question of whether all writing is to be regarded as phonographic (cf. DeFrancis 1989).

The basic distinction of phonographic vs. morphographic writing systems is based on the fact that only linguistic levels consisting of a (relatively) closed set of units can serve as base levels of writing systems. Phonemes, syllables, and morphemes are potential – and not always equally suited – candidates for the basic linguistic unit to be represented by graphemes. By contrast, words (in the sense of polymorphemic units), sentences, or texts are ill-suited for this purpose (cf. Meletis accepted b; Sampson 2015: 32). Consider a writing system whose graphemes correspond with sentences or whole texts (cf. the notion of *discourse writing* in Hill 1967), which would necessarily consist of an infinitely large inventory of graphemes and would thus strain the memory of its readers and writers to an unimaginable degree. This seems logically impossible.

Now that we have tackled the language system that underlies writing systems, let us look more closely at the modules that they are composed of. The first of these modules provides the writing system with its visual/graphic appearance; at the core of this module are so-called *scripts* as inventories of visual/graphic shapes that are referred to as *basic shapes*. However, not only basic shapes such as |R| or |a| in Roman script as well as digits, special characters, and punctuation marks constitute the visual substance of writing systems, which is enriched by an abundance of additional resources such as bold print, underlining, layout, etc. Therefore, this module cannot be reduced to scripts and is rather broadly termed GRAPHETICS. In Figure 1, it is intentionally positioned outside of the language system's boundaries as from the perspective of a given language, the choice of a script (and other visual resources) is, in theory, arbitrary.[24] Language could be written using various scripts – and myriad examples over the course of history reveal that they indeed have been. Consider as an extreme example Azeri, which has been

24 In fact, visual resources that are superimposed upon scripts (and thus often considered 'suprasegmental'), including bold print, italics, etc., appear to be more universal than the scripts themselves, as they are used across many writing systems.

written in a multitude of scripts: Arabic, then Roman, then Cyrillic, then, finally, again Roman (cf. Hatcher 2008; Section 8.1). German could also well be written with other basic shapes than those of Roman script, and indeed, it was, for instance in Fraktur. Inversely, scripts can be used to materialize a great number of writing systems: Take Roman script, which is obviously linguistically independent as it is employed by myriad writing systems; the same applies to Cyrillic script. In a nutshell, the link between a language and a script that is constituted by a given writing system is not fixed, as scripts can be switched and one and the same script can also be used in multiple writing systems. Coulmas (1996b: 1380, emphasis in original) provides a definition of *script* that fits the model proposed here: "*Script* refers to the actual shapes by which a writing system is visually instantiated. [...] Every writing needs for its materialization a script, but there is no necessary link between a particular script and a particular writing system".

Thus, prototypically, there is no link between a writing system's script and its underlying language. However, there are exceptions in which a connection can indeed be assumed. To explain this, it is vital to compare the processes of *script creation* and *script adoption* (cf. Rogers 2005: 4f.). Script creation refers to the rare instances in which, in the context of the development of a new writing system, a script is created from scratch. In this case, the basic graphetic units – the script's basic shapes – are closely linked to the linguistic units they are graphematically related to; since the script is designed specifically for a given language, they are 'custom-tailored'. This is reflected not only in the number of basic shapes that may in such cases roughly correspond to the number of linguistic units that need to be represented by the writing system but also in special properties such as pictography, which designates a special relationship between the graphetic substance and the linguistic content and is characteristic of the first writing systems ever invented (cf. Section 6.2). By contrast, both the number of basic shapes and pictography are of different weight in the process of script adoption, where an existing script is adopted and, in many cases, specifically adapted for the writing system of a language that it was *not* originally devised for.

A special comment shall be made about those first times that writing was (independently) conceived, which Daniels (2013: 56) calls *ancient grammatogenies*. Presently, following the hypothesis of the polygenesis of writing, it is assumed that this occurred at least[25] three times, uncontestably in Mesopotamia, China, and Mesoamerica. These ancient creations of writing systems as instances of the invention of the cultural technique of writing itself represent special cases, as here, right from the beginning, basic shapes were created in close connection with the linguistic units they stood for. As mentioned above, they were 'custom-tailored'. These are true instances of script creation. What distinguishes these ancient creations of scripts and writing systems from modern creations is that the inventors

25 Rogers (2005: 4) notes that "[s]ome scholars have claimed that the Egyptians and the people of the Indus Valley also invented writing" but that "these claims are controversial".

of the former did not have any examples to fall back on. They did not only invent writing systems, they invented writing itself. In *modern grammatogenies*, i.e. modern creations of writing systems – whether they are *sophisticated* or *unsophisticated*[26] – there is always at the very least the pre-existing knowledge of the concept of writing.

In terms of frequency, script adoption and adaptation can be declared the default. They are common strategies chosen in the creation of new writing systems, where pre-existing and, crucially, established scripts, most prominently Roman script, are adopted for economical, technological, as well as political and sociocultural reasons – all of which will be discussed in the context of the graphetic sociocultural fit (cf. Section 8.1). People working on the new writing system can either make use only of the existing basic shapes or, if necessary, refunctionalize or modify them, invent new shapes, or omit superfluous ones (cf. Daniels 2006a).

Notably, labeling this process merely *script* adoption is often inadequate, as it is most frequently not solely basic shapes that are borrowed from one writing system to another. What is adopted instead is basic shapes with their prototypical links to linguistic units, in other words: underspecified graphematic relations or underspecified *graphemes*. Therefore, it is terminologically more adequate to speak of *grapheme adoption*. For instance, when the basic shape |a| is adopted by a new writing system, naturally not in isolation but together with other basic shapes from Roman script (if not the entire Roman script), it will prototypically be employed to represent a vowel, more specifically an open unrounded vowel. While the linguistic units – in this case, phonemes – that correspond with the original grapheme and the borrowed grapheme, respectively, might not be identical (crucially, they are part of different language systems), they are often highly similar. Yet, there are also other – much rarer – cases of ‘pure’ script adoption. An example is the invention of the Cherokee writing system. In this system, a number of uppercase basic shapes have been transferred from Roman script. In Cherokee, however, they bear quite different graphematic relations than in other writing systems using Roman script, not least because the Cherokee writing system is syllabographic and not alphabetic, i.e. segmentally phonographic. Thus, for example, the basic shapes |A|,[27] |W|, and |L| represent the syllables /go/, /la/, and /tle/, respectively. This and only this scenario is to be considered as ‘pure’ script adoption, as units of scripts are transferred from one writing system to another entirely stripped of the graphematic values they were originally associated with.

As this example underlines, scripts can be analyzed divorced from the graphematic relations they take part in, and their link with specific languages

26 These are Daniels’ terms. In an *unsophisticated grammatogeny*, the inventor of a writing system “was not literate in any language but only knew by observation that writing existed” (cf. Daniels 2007: 56f.). By contrast, in a *sophisticated grammatogeny*, the inventor is literate (cf. also Daniels 1992: 85; Daniels 2013: 55).

27 Graphetic units such as basic shapes and graphs are enclosed in vertical strokes | |, while graphematic units such as graphemes are notated in angle brackets < > (cf. Berg & Evertz 2018: 190).

is rather arbitrary. This justifies their position outside of language. However, they must imperatively be placed within writing systems, as at any given point in time, a writing system must be materialized by a script, or, in the case of *biscriptality* or *multiscriptality*, by more than one script (cf. Section 8.1). As mentioned above, however, the script as the visual manifestation of a writing system can be changed when writing systems switch from one script to another. This famously occurred when a switch from Arabic to Roman script was mandated for Turkish in 1928 (cf. Wood 1929).[28] Yet, despite their interchangeability, scripts are indispensable parts of writing systems; they take on the important role of providing the material makeup for graphemes, and indeed, given the salience of a writing system's visual appearance, users perceive scripts as intricately linked with the linguistic units they signify, often equating them *pars pro toto* not only for their writing system but for their language. Thus, while theoretically, scripts are not linked to a specific language, users' reality is different as they strongly associate scripts with language, rendering politically motivated switches of scripts invasive disruptions and delicate affairs.

Another reason that justifies positioning the graphetic module outside of the language system is that scripts can also be used for non-linguistic notational purposes. Writing, as defined above, is interpreted as the graphic representation of language; as such, it is a special form of notation, the notation of language. Not only language can be noted down: take the notation systems of mathematics, dance, or music, in which scripts or parts of scripts are utilized for purposes other than writing. Notably, these non-linguistic functions are only secondary functions of scripts which are indeed primarily used for writing (except if they have been devised specifically for a special purpose).

Crucially, the central reason that warrants and even requires positing the independence of the graphetic module from language is that it is, in fact, the materiality of writing that lends writing its idiosyncratic features – features that cannot be explained with recourse to language (or speech as the dominant modality of language) and do not necessarily have a parallel in it. For example, the relevant dimension for writing is space, not time. The spatial arrangement of

---

**28** Note that technically, the Turkish writing system that employs Roman script is a different writing system than the Turkish writing system using Arabic script. In this case, not only the visual substance (the basic shapes) of graphemes is switched but also their graphematic relations. In other words, the two systems differ typologically: the first system is an alphabet, the latter an abjad. This situation in which two scripts used for the same language constitute different writing systems is referred to as *intersystemic biscriptality* (cf. Section 8.1). By contrast, hypothetically, two different scripts can also materialize one and the same writing system if only the basic shapes are switched out while the linguistic units they correspond with remain stable (this is more or less the case for biscriptual Serbian, which is written in either Roman or Cyrillic script). Most frequently, however, as in the case of Turkish, when a language has been/is written with two or more scripts, the different graphematic relations associated with those scripts mark them as distinct writing systems, even if the underlying language system is the same.

units of writing, their pre-segmented nature, i.e. the fact that they, unlike 'units' of speech which must be extracted from a continuum, are by default segmental, the permanence of writing – all of these features (and more) stem from the substance of writing, its graphetics. They are not dependent on speech or any linguistic subsystem (morphology, syntax, etc.). Consequently, these features must take center stage in the investigation of graphetics as a part of a theory of writing.

Next is the core of writing systems as semiotic systems: GRAPHEM(AT)ICS[29] fulfills the central task of relating graphetic units to linguistic units. Accordingly, the graphematic module is constituted by relations, which can reasonably be interpreted as semiotic relations.[30] The smallest of these relations is simultaneously the smallest linguistically functional unit of writing systems, the *grapheme*.[31] Since the grapheme is an infamous unit let alone a controversial term, it will be discussed in detail in Section 2.2. In a nutshell, in the present model, the grapheme is conceived of as a dyadic sign in the Saussurean sense, consisting of two inseparable constituents: the *signans* (pl. *signantia*), a graphetic unit (the most elementary of which is the basic shape), and the *signatum* (pl. *signata*), a linguistic unit such as a phoneme or a morpheme. A factor that makes identification of graphemes in a given writing system challenging is that semiotic relations are not biunique and stable, meaning that there is variation with respect to both constituents: on the one hand, there can be multiple signantia in the sense of abstract basic shapes – such as |a| and |ɑ| which differ visually but have the same graphematic value – or concrete graphs – possibly infinite concrete physical manifestations of basic shapes, such as |a|, |a|, and |a|, which materialize the same basic shape in different typefaces. The situation is even more complex for distinct basic shapes such as |v| and |f| which can have the same signata: in the writing system of German, they are both graphematically related to the phoneme /f/. Vice versa, there can also be multiple signata and a single basic shape such as the aforementioned |v| can participate in more than one graphematic relation: in the German writing system, |v| is not only graphematically related to /f/ but also to /v/ (cf. also Section 2.3 on allography).

---

**29** In most works, *graphemics* and *graphematics* are treated as synonyms (cf. Glück 2016e: 253). Some scholars, however, distinguish between them: Fuhrhop & Peters (2013: 203), for instance, interpret the adjective *graphemic* to be related to the grapheme as a unit, while *graphematic* refers to graphematics as a module of writing systems and, in turn, as a subpart of the grammar of languages equipped with a written modality.

**30** The treatment of writing systems as semiotic systems in which written units are signs of linguistic units is neither uncontroversial nor unproblematic (cf. Harris 1994). However, in line with a number of recent approaches (cf. Klinkenberg & Polis 2018; Rizza 2018), I opt for a semiotic analysis. In doing this, I do not claim to be able to provide an answer to the question of *how* written units signify units of language (cf. Harris 1994: 45), as I merely posit that the two are semiotically related. Nevertheless, the semiotic parameters presented in Chapter 6 will help to characterize the nature of the relations between them.

**31** Note that there are some exceptions to this. In some writing systems such as Chinese, there are subsegmental graphematic relations that are smaller than graphemes (cf. Section 2.2). In most cases, however, *grapheme* and *smallest graphematic relation* can be considered synonymous.

The complexity of these relations is captured by transparency (cf. Section 6.4) and uniformity (cf. Section 6.5), two parameters of the linguistic fit, i.e. the question of how well a writing system suits its language.

As was mentioned above in the context of the language system underlying the writing system, there is commonly one linguistic level that the graphemes predominantly relate to. However, this level is not necessarily fixed, and thus, signata of a writing system's graphemes need not consistently be phonemes, syllables, or morphemes. Consider the writing system of Japanese, in which multiple scripts are in use simultaneously: *kanji* are used morphographically, while the two so-called *kana* scripts (*hiragana* and *katakana*) are used syllabographically, i.e., in a broader sense, phonographically. Additionally, there is *romaji*, Roman script that is used alphabetically. This kind of systematic type mixing was also characteristic of several ancient writing systems, including Egyptian hieroglyphics, where morphographic and phonographic components were mixed. Notably, also in relatively 'pure' phonographic systems such as English, we find elements such as <&> or <§> which are used morpho-/logographically. Crucially, though, they represent 'exceptions' as they are few in number and fulfill special functions. This distinguishes the type mixing in Japanese, where it is central to the system in that it concerns graphemes that fulfill prototypical linguistic functions, from that in English, where non-phonographic graphemes are a peripheral phenomenon.

In a nutshell, the graphematic module contributes relations to a writing system, functioning as the vital link between a language system and a writing system's graphetic module. It generates graphematic units such as graphemes but also larger units such as graphematic words, among others. Graphematics is central for an explanatory grapholinguistic theory as it enables an evaluation of the linguistic fit of writing systems (cf. Chapter 6). Questions relevant in this context are: Are there enough basic shapes to transparently correspond with all linguistic units, e.g. with all phonemes or syllables of a language? In general, what is the relation between basic shapes and linguistic units? Is there a one-to-one relation or are there one-to-many relations as sketched above? In most of the world's writing systems, the latter prevail. Given these multi-basic shape-to-linguistic unit correspondences (and, inversely, basic shape-to-multi-linguistic unit correspondences), not to mention other relevant aspects such as (in some writing systems) capitalization, word division, etc., there often exist many possibilities to write a given word. Notably, all of these possibilities must conform to a system's *graphotactics*, rules that restrict the combination, position, etc. of written units (cf. Section 2.4). For example, in German, the word *Fuchs* 'fox' could be written *<Fux>, *<Fuks>, *<Fugs>, *<Vux>, <Fuchs>, not to mention the corresponding uncapitalized variants. All of these variants are possible according to the graphematics of the German writing system; in other words, they are located within the so-called *graphematic solution space* of the word *Fuchs* (cf. Neef 2005, 2015). In general, this space is defined as the sum of possible variants – i.e. those licensed by the writing system – of graphematically representing a given word. However, as indicated by the asterisks, most of these possibilities are deemed 'incorrect'. Although many of them would be understood by readers and would, thus, successfully fulfill their communicative purpose, they are considered incorrect from the perspective of a standardized norm – an orthography. Descriptively, we deal with the categories possible vs. impossible within the system, while prescriptively, we deal with the normative categories of correct vs. incorrect.

This leads us to the final module of writing systems: ORTHOGRAPHY. As the statement that orthography is only a module of writing systems already implies, these two terms cannot act as synonyms. In other words: *orthography* and *writing system* shall not be used interchangeably. In short, writing systems are entire systems consisting of the modules presented here. As phenomena, they are theoretical reconstructions of a system that is constituted by the empirical sum of regularities present in users' actual use of the written modality in a given language. As such, writing systems are not *per se* prescriptive as they do not single out correct 'spellings' – a word already carrying the connotation of norms – for words and larger units such as sentences. They do this only if they are equipped with an orthographic module. Crucially, thus, what distinguishes orthography from the other modules of writing systems is that it is optional. If a writing system is equipped with an orthographic module, as illustrated in Figure 1, it superimposes the graphematic module. By codifying certain variants as correct – via rules in rulebooks, dictionaries, etc. – orthography restricts the possibilities of a writing system. Often, authorities of linguistic policy such as language academies, ministries of education and culture, etc. make decisions about what is codified as orthographically correct. These decisions may, of course, in the best-case scenario, be based on the graphematics of a writing system, more specifically on the scribal practices and implicit conventions of its users. Orthographies for which this is done are indirectly shaped by an invisible hand (cf. Keller 2014): in short, this means that users' choices influence the writing system, whereby they can indirectly affect what is taken over into and codified by the orthography. This is, however, as mentioned, frequently decided by authorities involved in linguistic policy. Frequently, but not always: for English, which is a self-organizing orthography, no such official authority – such as the *Council for German Orthography* for German – exists. Yet, publishers of dictionaries wield some power as they decide which entries to include in dictionaries, ultimately shaping what users perceive as orthography.

Since orthography is optional, there exist writing systems without it. In fact, diachronically, it is a relatively recent development: many systems that are now equipped with a codified orthography have previously done without one. Note that communication in systems without an orthography does not automatically pose a problem. Depending on the level of transparency and uniformity of a writing system's graphematic relations, the graphematic solution space, i.e. the sum of possible spellings for the same utterance, might be either small or large. Take as an example Finnish, where graphematic relations are almost biunique, which renders the graphematic solution space small and means that the orthographic module does not need to select one variant among many possible variants. Yet, even if graphematic solution spaces for words are large, the lack of an orthography might not be fatal: as long as the message that the sender wants to convey reaches the addressee relatively unscathed, communication is successful and the writing system serves its most crucial function. However, from a different perspective, the orthographic module offers advantages for communication: if it is systematic in that it makes use of regularities that already exist in the system and, for example, generalizes them, this external systematization can help make a system easier to use for readers as well as writers. This function of orthographies will be discussed in the context of their systematic fit (cf. Chapter 5).

The difference between graphematics as the constitutive module of writing systems and orthography as an optional module can be summed up as follows: graphematics encapsulates both everything that a writing system allows writers to do to communicate (or do other things they intend to do) through written language, i.e. the sum of well-formed possibilities, and everything that users actually do, i.e. the sum of all observable empirical regularities. It is the latter that is at the center of graphematics. Orthography, by contrast, standardizes – and simultaneously curtails – the possibilities and resources offered by the graphetic and graphematic modules of a writing system. Depending on how orthography interacts with these modules, it can be useful or present a complication for users. Relevant questions that are raised by the optional nature of orthographies are: Which writing systems have orthographies, which systems lack them, and what are differences between these two types of systems?

To conclude the presentation of the multimodular model, I want to mention that not only the graphematic module can be standardized by an orthographic regulation, as is implied in Figure 1. The graphetic module can also be subject to norms. However, norms pertaining to the graphetic module are, unlike norms concerning the graphematic module, commonly not externally codified (with exceptions, such as stroke order in Chinese, cf. Section 7.3); consequently, they are neither palpable nor binding in the same way codified rules are perceived to be. The 'ortho-graphetics' that I am referring to rather exists (or does not exist, for that matter) implicitly in the writers' and readers' competence. For instance, a mostly implicit convention that is, however, sometimes also explicated, dictates:[32] do not use flashy or playful typefaces or, more generally, any typeface that could be perceived as inadequate when designing a job application. This depends on the job, of course, as extravagant typefaces can be adequate and even expected for job applications in the fields of design, advertisement, etc. A very demonstrative example of what is perceived as an 'ortho-graphetic' mistake or better 'misstep' is the widespread use of the typeface Comic Sans – for all imaginable purposes. What originated as a typeface designed for children and is still largely perceived as a playful typeface adequate for, among other things, invitations to children's birthday parties, is now being used, to name only a few contexts, on gravestones, in

32 The fact that this convention is sometimes explicated (in guides, etc.) implies that not everyone is expected to have the necessary implicit knowledge about graphetic conventions or rules.

medical reports, in information brochures for rape victims, etc. These uses of the typeface and similar graphetic behavior, while resisting to be categorized as traditional 'mistakes' due to the lack of a codified rule that is broken, are still perceived as mistakes by other users of the writing system and in some cases even sanctioned by a portion of the members of a literate community – such as when an HR manager throws out a job application because it uses Comic Sans (cf. Meletis 2020a).

While this introduction to the book's descriptive part served as an overview of the general structure of writing systems, the following three chapters will delve deeper into the three modules of graphetics (Chapter 1), graphematics (Chapter 2), and orthography (Chapter 3). Note that while this part of the book is undeniably structuralist (or at least heavily influenced by structuralism), it aims to eschew the practice of rigidly adhering to the rules of the structuralist paradigm (which is often done since as a subject, writing was 'late to the structuralist party', cf. Schroeder 1981: 132). Although they can of course also be descriptive, non-structuralist psycholinguistic and sociolinguistic perspectives are integrated into the theory in the explanatory part of this book (Part III) since, arguably, they can offer answers to *why* writing systems are structured the way they are, a question that a purely structuralist analysis cannot (and does not want to) answer.

# 1 Graphetics

As a grapholinguistic subfield, graphetics studies the materiality of writing, covering all phenomena pertaining to the graphetic module of writing systems.[33] It is not only the material auxiliary discipline to graphematics but also approaches questions that pertain not primarily or exclusively to linguistics but rather to a number of other neighboring disciplines such as philosophy, didactics, neuropsychology, and art history. Accordingly, graphetics can broadly be defined as an interdisciplinary field of research. Unsurprisingly, works that have treated graphetic questions are scattered across different disciplines and there is a palpable lack of reception beyond disciplinary boundaries (cf. Spitzmüller 2016: 103).

It is symptomatic that the term *graphetics* is absent from much of the literature on writing, let alone linguistic literature in general (cf. Rezec 2009: 8). This is striking insofar as the graphetic module of writing systems is just as relevant as the graphematic and orthographic modules. In the end, one cannot write or read if there is no visual (and/or tactile) substance. Ignoring this fact would do the study of writing injustice. Several recent studies prove that an investigation of the interplay between the graphetic and the graphematic modules is a promising endeavor as they discovered striking correlations between graphetic form and graphematic function (cf. Primus 2004; Bredel 2008; Fuhrhop, Buchmann & Berg 2011). In any case, a deeper and more fine-grained analysis of the structure and use of scripts and other visual resources employed in writing can, even if it is not located within the immediate core of linguistics, only enrich grapholinguistic research.

Graphetics is characterized by an often-drawn analogy with phonetics, which is unsurprising given that the term *graphetics* was coined based on *phonetics*. Following this view, graphetics is to graphematics what phonetics is to phonology and graphetics is to writing what phonetics is to speech. Like phonetics, graphetics studies language and is thus inherently linguistic. However, it does study questions and uses methods that are in the periphery of what is considered linguistic, leading some to argue that they are in fact not linguistic.[34] Thus, echo-

---

33 Large parts of this chapter overlap with Meletis (accepted a).

34 An example of such a question is the perception of different typefaces. Not only can the physiological aspect of the perception of different typefaces be compared to answer questions such as *Which typeface is more legible?*, but due to the often connotative nature of typefaces (or handwriting), the emotional response to them can also be studied. Take the study by Velasco et al. (2015) who instructed participants to match round or angular typefaces with taste words and found that round typefaces are associated with attributes such as “sweet” while angular typefaces are associated with

ing a distinction made in phonetics (cf. Ladefoged 1997; Laver 2017), a linguistic graphetics could be differentiated from a broader general graphetics. This is a question I will leave open for future discussion. Furthermore, it is paramount to note that graphetics is certainly not only an auxiliary discipline to graphematics. Graphetics and graphematics certainly go together, although their relationship is not quite symmetrical: while it is possible to conduct graphetic research without being interested in graphematic matters, the opposite cannot be posited. Just like we usually do not do phonology without phonetics (at least not completely), why should we do graphematics without graphetics? Without graphetics, writing would be invisible and/or intangible[35] – it simply would not exist. Granted, there is some truth to what those linguists and semioticians who disregard graphetics claim, the most famous of whom is probably Ferdinand de Saussure (1916: 143): for the meaning of an utterance, it often does not matter how writing appears, "an A is an A is an A" (Stöckl 2004: 5f., my translation; cf. also Assmann 1988: 144) no matter what typeface is used or what an individual's handwriting looks like exactly (but cf. for the connotative relevance of its appearance below). However, no one can deny that it is imperative that writing 'looks' at all, i.e. that it is materialized in the first place. In any case, the materiality of writing is not just an accidental side issue but constitutive of writing.

Even though, as noted, some parallels exist, the analogy phonetics/graphetics is also the basis of several misconceptions. One of them is that graphetics studies materiality in a solely formal manner and is not concerned with functions (cf. Spitzmüller's discussion and criticism of a structuralist two-world ontology, cf. Spitzmüller 2013: 124; Krämer 2001: 95–105). This characterization falls short: graphetics *is* also interested in functions, specifically the functions of the written substance itself (and the practices involved in their production and perception) rather than the functions of the linguistic information visualized by that substance. In the analysis of a given product of writing, for example, graphetics does not concern

"bitter", "salty", and "sour". The authors hypothesize that this could be caused by the fact that round typefaces are easier to process. This study is undeniably graphetic; it is debatable, however, to which degree it is linguistic.

35 The word 'intangible' is included here because this also holds for braille writing, which works (primarily) on a tactile rather than a visual level. Although sometimes, in restrictive conceptions of writing, braille is not regarded as a form of writing (or simply not mentioned as such, cf. Glück 2016c: 593), it is a graphic representation of language and it should be counted as writing. 'Graphic', which derives etymologically from Greek *γράφω gráphō* 'scratch, carve' emphasizes this broader reading; this, however, should not obscure the difference between *tactile* vs. *visual*, which is crucial. Yet, since embossed marks as well as visual marks are *material* (and visual marks are always also in a way tactile and vice versa), they are both studied by graphetics, which is with good reason defined as the study of the *materiality* of writing rather than the study of the visuality of writing. As Spitzmüller (2016) notes, braille writing proves that writing does not necessarily have to be constituted visually. For that same reason, Harris (2005) proposes that the feature of *spatiality* rather than *visuality* is constitutive of writing.

itself with denotative meaning[36] but with the connotations that are evoked by visual features such as color, typeface, type size, highlighting such as bold print, italics, etc., and with the question whether an additional[37] layer of meaning – sometimes the crucial one – is constituted by the visual appearance of a written utterance. Consider, for example, *pseudoscripts* or *typographic mimicry* (or *exotypes*, cf. Alessandrini 1979; Haralambous 2007: 414), terms denoting that a typeface is designed to imitate the look of a different script (cf. Coulmas 2014: 16–19). In the examples in Figure 2, typefaces in Roman script are made to resemble Devanāgarī, Chinese, and Arabic, which is achieved solely by the respective type design (and, importantly, only works if readers have the necessary graphetic knowledge). The words themselves could have been written in a prototypical typeface of Roman script, of course, in which case, however, the specific cultural meaning evoked by the culturally specific type design would have been lost. This cultural meaning is a fundamentally graphetic matter.


Figure 2. Pseudoscripts

Some have criticized the term *graphetics* and what it connotes due to the terminological analogy with *phonetics*, claiming that these two disciplines cannot be read-

36 A possible graphetic question that concerns the denotative meaning is: To what degree must graphs differ in order to be perceived and categorized as materializations of distinct basic shapes instead of as two materializations (i.e. allographs, cf. Section 2.3) of the same basic shape? Categorical perception at this level is a solely visual matter. However, even if the graphs differ visually to such a degree that they are in fact members of two basic shapes, the question is if one can speak of a different 'denotative meaning' since at the graphetic level, we are not concerned with the linguistic units that basic shapes relate to. In fact, the assignment of basic shapes to graphemes and thus, their correspondence with linguistic units, is a matter of graphematics, not graphetics. For example, that in writing systems using Roman script (take German and English as examples), the visually similar but still distinct |g| and |ɡ| belong to the same grapheme cannot be decided on visual grounds (at least not solely), which is more obvious for the visually dissimilar shapes |σ| and |ς| which are allographs of the grapheme <σ/ς> in the writing system of Greek.

37 The treatment of these functions as *additional* functions and an *additional* layer of meaning – i.e. connotative meaning – is criticized by Ludwig (2007), as he argues that this classification as 'surplus', as something secondary to linguistic denotative meaning, hinders a systematic distinction between linguistic functions and visual (or, more generally, material) functions that are performed by written utterances or their production and perception. However, *graphetics* as proposed here aims to systematically investigate the functions of visual materiality independent of linguistic functions.

ily compared. One such critic is Konrad Ehlich (2001); he proposes an alternative designation, *transindividual graphology*. Transindividual is self-explanatory, and indeed, graphetics is necessarily trans-individual as it does not primarily study the writing of individuals but of, for example, entire literate communities. However, the second part of his proposed term, the polysemous *graphology*, needs to be commented on. Firstly, this term proves problematic because it has already been used by a quite different field that Ehlich seeks no association with, a field that attempts to reconstruct psychological profiles of writers based on (visual) features of their handwriting (cf. Paul-Mengelberg 1996). While a descriptive analysis of writing's visual features is of course not *per se* problematic, it is the association with psychological traits that has been overwhelmingly criticized as being unscientific (cf. Dürscheid 2016: 219f.).[38] Ehlich, however, aims to reappropriate the term *graphology*; he claims that its suffix *-logy* as found in designations of other scientific disciplines and linguistic subbranches, which also establishes a direct terminological parallel to phonology, highlights that the material subsystem of writing has its own systematicity. What Ehlich means by 'systematicity' is the fact that writing is spatially organized in a way that allows studying it as a visual system completely without the consideration of linguistic facts. This is not the case in phonetics, where the meaningful organization of sounds is not studied, as this would already be a matter of phonology. This lack of systematicity in phonetics is what makes the analogous term *graphetics* unsuitable for writing, Ehlich (2001: 65, emphasis in original) argues:

> What is termed graphetics [...] should be conceived of as [...] *transindividual graphology* in the same sense in which phonemics (or phonology) is used: the scope of analysis [...] is to come to a *theory of scriptural form*, – i.e., its purpose is to reconstruct how, to which extent, in which ways and to which results the optical, physiological and psychological possibilities are made use of in order to establish a writing system [...]. In the center of interest [...] are the description and analysis of functionability and functionalizing of the objects of graphetics for establishing scriptural structure. This structure is a systematic phenomenon of its own type.

Due to the difference in medium (acoustic vs. visual), the dimension that is relevant for speech, and thus, phonetics, is time, while for writing and graphetics, it is (primarily)[39] space (cf. Dürscheid 2016: 32f.). The terminological analogy between the two terms, thus, in any case, works at a very abstract level, implying only that what is studied by both disciplines is the etic level, i.e. materiality;

---

38 Graphology must be distinguished from *forensic handwriting examination*, which is concerned with testing the authenticity of handwritten texts, identifying the (hand)-writer of texts, and determining the conditions under which a text was produced (cf. Michel 1996: 1036; Fuhrhop & Peters 2013: 185; Harralson 2013).

39 Note that from the dynamic perspective of production (and perception, for that matter), time does play a role also for writing, as writing and reading processes are of course bound to time. However, from the perspective of the product, i.e. the written text, time is not relevant. It is, by contrast, relevant for the product(s) of speech, i.e. spoken utterances.

this does not, however, preclude that the material level has an internal systematic structure in writing.

As it disregards the linguistic level, Ehlich's proposed transindividual graphology would not supplant but rather complement graphematics, i.e. the grapholinguistic subfield that deals with precisely the linguistic aspects of writing. While I wholeheartedly agree with Ehlich that there is a spatially-based systematicity to the materiality of writing that speech is lacking, I do not agree that the term *graphetics*, on the grounds of its analogy with phonetics (rather than phonology), conceals this fact. In fact, the inner systematics of the graphetic module will be the subject of Section 1.2 below.

## 1.1 Subbranches

Analogous to a subdivision of phonetics, three graphetic subdisciplines can be postulated (cf. Fuhrhop & Peters 2013: 182–183). They result from a very simplified model of communication that starts with production. *Productional graphetics* asks questions that pertain to the material aspects of the writing process. On the one hand, it focuses on the cognitively lower and unconscious levels of writing: which fundamental processes are involved in producing sequences of graphs in handwriting?[40] To consider also modern technologies: which processes are involved when typing on a keyboard or swiping on a touchscreen? These questions are primarily of physiological and psycholinguistic nature. An example of applied productional (and perceptual) graphetic research is the study of so-called *character amnesia* in Chinese and Japanese (cf. Xu 2015). It designates the phenomenon that users of these systems forget how to handwrite specific morphographic graphemes that they could formerly write (cf. Section 7.1.2). Interestingly, in many cases, users are still able to read these graphemes, underlining that reading and writing processes are to some degree independent of one another.[41] On the other hand, choices located at higher and conscious levels of production and yet

40 Movements in handwriting are studied by a field called *graphonomics.* This term was coined in the 1980's and defines a "multidisciplinary emerging field focused on handwriting and drawing movements" that has made an "important contribution to the field of motor behavior by developing models aimed to conceptualize the production of fine motor movements using graphical tools" (van Gemmert & Contreras-Vidal 2015: 165). Because *graphonomics* also concerns itself with the production of non-linguistic graphic material, it is not a subfield of graphetics, although there is certainly a great deal of overlap between graphonomics and graphetics.

41 Another striking example of this is *pure alexia*, also referred to as *alexia without agraphia.* People who suffer from this condition have lost their reading abilities, while visual recognition in general and writing skills are preserved (cf. Rupareliya, Naqvi & Hejazi 2017). Hence, a person can write something, but even immediately after, the person is not able to read what they have just produced themselves – a reflection that in the brain, regions that are responsible for reading can be impaired while regions for writing remain unaffected.

concerned with visual aspects are also studied by productional graphetics: from a sociolinguistic perspective, for example, questions can be asked about writers' motivation to choose a specific typeface or a specific form of highlighting (bold instead of italics or underlining, etc.). Choices at all levels of writing, including the material, are – to some degree – "acts of identity" (cf. Hatcher 2008), whether they are conscious or unconscious. This means that these choices refer indexically and sociosemiotically to producers and facets of their (self-constructed) identity. The following questions are relevant in this context: what was the writer's intention in designing a text in a specific way, and was it motivated socioculturally – if so, how? Does a text's producer want its graphetics to convey belonging to or distance from a certain social group? Ultimately, all questions that are asked in graphematics can be studied here as well – only at another level.

The second subbranch of graphetics is likely the most 'traditionally' linguistic one in that it is solely descriptive. *Script-graphetics* or *descriptive graphetics* (from German *Skriptgraphetik*, cf. Meletis 2015: 42f.; Fuhrhop & Peters 2013: 183) visually analyzes products of writing divorced from the processes of production and perception. This, however, does not imply that a descriptive analysis cannot occasionally spawn questions pertaining to other graphetic subdisciplines as well, for example, how production and the involved surfaces and instruments have affected the visual shape of a product of writing. This question of why writing appears the way it does in a final product is indeed of relevance. A demonstrative example comes in the form of the visual appearance of an entire script that has developed over a longer period of time: the Burmese script, which, in Burmese, is also referred to as *calonh* 'round script' (cf. Coulmas 1996a: 55; Watkins 2009: 170; cf. Figure 3). It is predominantly curved because it was traditionally written on palm leaves whose fibers are linear; this means the production of angular basic shapes would have caused them to rip. Regarding these issues, descriptive graphetics is similar to neighboring, predominantly historically-oriented disciplines such as paleography and epigraphy. These can be seen as specialized subdisciplines of descriptive graphetics. The different levels of graphetics that are presented below in the context of the cartography of the writing surface (see next section) are based on a description of the spatial arrangement of writing and, thus, result themselves from a script-graphetic analysis.

မြန်မာစာရေးသားရာတွင် အပိုဒ် (၂၇) ပါ ဗျည်း (၃၃) လုံး
စာပေတွင် သုံးသည့် ဥက္ကလေးကိုလည်း အသုံးပြုလျက်ရှိ
စာပေတွင် စဝဂ်၌ စ၊ ဆ၊ ဇ၊ ဈ၊ ဉ ဟု ဥက္ကလေးနှင့်သာ ရှိ
ကြီးမှာ ပါဠိတွင် ဥက္ကလေးနှစ်လုံးဆင့်သည့် ပါဠိဆင့်စာလုံ
သို့သော် မြန်မာစာတွင် ပုဂံခေတ်ကစ၍ စဝဂ်ကို စ၊ ဆ၊ ဇ

FIGURE 3. Extract from the Burmese Wikipedia page covering the Burmese writing system

The third and final subbranch, and arguably the most prominent of the three, is *perceptual graphetics* (cf. Meletis 2015: Chapter 4). Like productional graphetics, it is not predominantly a linguistic subfield but rather one that is enriched by research from fields such as psychology, the cognitive sciences, and neurobiology. It is concerned mainly with the processes of perception, recognition, and – at the highest level – reading.[42] How is a basic shape or a word that is itself made up of a sequence of basic shapes/graphemes recognized? At a higher – but not necessarily conscious – level, sociolinguistic questions can be asked that are symmetrical to the questions studied by productional graphetics: which emotions are evoked in the perception of different typefaces? Which connotations do typefaces carry? What is the attitude towards a specific style of writing (a specific typeface, handwriting)? In general, what knowledge do users have about graphetic practices and what do they think about them, i.e. what are (their) 'graphetic ideologies'? A striking example of the importance and the reality of a sociolinguistic perceptual graphetics is the passionate discourse about the dislike for the typeface Comic Sans, especially online (cf. Meletis 2020a). This is largely a sociolinguistic issue, but since it has at its core the materiality of writing, it is also a matter of graphetics.

As evident from the questions asked by these graphetic subbranches, there exist, as in phonetics, two methodological strands or perspectives which Günther (1990b) terms *symbol graphetics* and *signal graphetics* (cf. also Bredel 2008: 24). Symbol graphetics describes and attempts to categorize the graphetic resources of the world's writing systems and often calls on extra-graphetic, i.e. graphematic information in order to assemble graphetic categories (such as basic shapes, see below) and discover universals or universal tendencies (cf., for example, the studies by Changizi & Shimojo 2005 on the number and complexity of elementary forms in the basic shapes of the world's scripts or Morin 2018 on the predominant cardinality of these elementary forms, cf. also Section 7.1.3). As such, symbol graphetics bundles questions from disciplines such as linguistics, cultural studies, philosophy, history, etc. The sociolinguistic questions listed above are examples of questions studied by symbol graphetics. By contrast, signal graphetics employs experimental methods borrowed from the sciences: the materiality of writing is studied divorced from the linguistic structures it is associated with, and what is of concern is optical stimuli and motor and perceptual processes involved in processing them, which are addressed using a range of methods, including eye movement studies, imaging technology, and graphonomics (see above). Accordingly, signal graphetics bundles graphetic questions from psychology, physics, medicine, IT, etc. The psycholinguistic questions raised above, thus, are largely of signal graphetic nature.

---

42 Reading, of course, already involves the linguistic level, and as such, reading processes cannot be treated solely by perceptual graphetics. For the study of reading, graphetic, graphematic, and psychological questions merge to what, essentially, is psycholinguistic research. What I want to underline here is the specific contribution that *perceptual graphetics* makes to this research by studying the material aspects of reading processes, aspects which are often ignored.

## 1.2 Cartography of the surface: Graphetic levels and units

The levels and units that will be presented in the following sections are constituted visually by “spaces of nothing” between them. These spaces will be referred to as *empty spaces*. As visual units independent of the linguistic information that they may materialize, graphetic units are to some degree universal but differ across writing systems based on where empty spaces are located. Accordingly, the *empty space criterion* states that graphetic (and, in turn, graphematic) units are constituted by empty spaces. It is based fundamentally on the gestalt theoretical principle of figure—ground that establishes “syntagmatic contrasts between a more important foreground or figure and a less important background” (Dressler & Kilani-Schoch 2016: 365; cf. Section 6.8), with visible graphetic material being the figure and empty space being its ground. A crucial theoretical question concerns the genesis of graphetic units and whether, in some cases, they are graphetic only secondarily, i.e. when empty spaces are determined by linguistic units. Note that the opposite could also be true, i.e. that graphetic units constitute linguistic units in the first place through visualizing them. This is the graphetic/graphematic chicken-and-egg-problem that Spitzmüller (2016: 108, my translation) addresses when asking “whether the text form merely makes visible an already existing informational structure or whether it itself creates its own informational structures”.[43]

The smallest empty space in the graphetic modules of most writing systems is the one located between basic shapes (see below for a definition), as can be evidenced in Roman script provided it is materialized in a typeface with spaces or spaced handwriting and not in cursive handwriting or a decorative typeface in which graphs are connected with each other. In Arabic, by contrast, even in print, there is no empty space between most of the segmental basic shapes as these are connected to each other. As illustrated schematically in Figure 4, different types of empty spaces constitute different subspaces of the writing surface. Spaces that occur universally are the *segmental space,* the *linear space,* the *areal space*, and the *holistic space* (cf. Bredel 2008, 2011; Meletis 2015: 115). These spaces are studied by *micrographetics, mesographetics,* and *macrographetics*.[44] Following Reißig (2015), I term the practice of spatially dividing the writing surface into subspaces of different hierarchical levels *cartography*. Crucially, the concatenation or combination of spaces from a lower level constitutes spaces at a higher level. This means that the

43 „[…] ob die Textgestalt lediglich eine bereits vorhandene Informationsstruktur von Texten *sichtbar* macht oder ob sie selbst eigene Informationsstrukturen *schafft*“ (emphasis in original).

44 These terms are adaptations of Stöckl’s (2004) typographic terminology (*micro-*, *meso-*, *macro-*, and *paratypography*). By substituting ‘typography’ with ‘graphetics’ (cf. Meletis 2015: 119), the terms are broadened, which reflects that typography is a part of graphetics. Typography is concerned with the printed – and nowadays, digital – word. A separate scribal practice and, in turn, graphetic subfield, is chirography, i.e. handwriting. Both fields – typography and chirography – are dealt with in graphetics.

*strict layer hypothesis*, originally formulated in phonology,[45] applies to graphetics as well. Accordingly, every holistic space is necessarily made up of areal spaces, which are made up of linear spaces, which are made up of segmental spaces.


Figure 4. Cartography of the writing surface: empty spaces and the graphetic levels and units they constitute, adapted from Meletis (2015: 116)

## 1.2.1 Micrographetics: elementary forms, graphs, basic shapes

The smallest space on the writing surface in which a graphetic unit is produced is the *segmental space*. This space and all the questions pertaining to it are studied by *micrographetics*. The central units at this level are the abstract *basic shape* and its concrete realization, the *graph*. Commonly, each basic shape fills its own segmental space. This marks one of the central differences between speech and writing: in writing, utterances are inherently segmented. Thus, what readers perceive are units that are made discrete by the empty spaces between them. In speech, by contrast, segmentation is a sophisticated task. There is a lively debate surrounding the claim that what is perceptually salient in spoken language is actually neither segments nor (phonological) words but syllables (cf. for a summary of this discussion Massaro 2011; Daniels 1992 discusses the relevance of this claim for writing, cf. Section 7.2.1). Of course, at the graphematic level, a single basic shape that occupies a

45 The original formulation of the hypothesis reads as follows: “We have proposed that a category of level *i* in the hierarchy immediately dominates a (sequence of) categories of level *i–1*” (Selkirk 1984: 26, emphasis in original).

segmental space can be in a graphematic relation with a phonological syllable, as in the syllabaries of the Japanese writing system: here, segmental graphetic and, in turn, segmental graphematic units relate to larger, non-segmental phonological units: the segmental <ぬ>, for example, corresponds graphematically to the mora /nu/. Notably, for an analysis of (only) the graphetic module, these graphematic relations are not of concern. To summarize, the fundamental perceptual difference between speech and writing is the fact that the most salient visual unit is segmental (with exceptions like cursive handwriting or Arabic script),[46] whereas the most salient acoustic unit is arguably non-segmental.

Basic shapes, the smallest graphetic units, are commonly complex, as they are made up of several segments.[47] In German-language grapholinguistics, these segments have been referred to as *elementary forms* in the past (originally German *Elementarformen*, cf. Berkemeier 1997: 242; Butt & Eisenberg 1990: 36; Meletis 2015: 65f.). They are the subject of controversial debate, as some scholars of writing opt to treat basic shapes holistically, i.e. not to break them down into smaller parts. The reason for this is the claim that segmentation is not of value, at least not from a graphematic perspective (cf. Neef 2005; Rezec 2009: 81; Wehde 2000: 74; Brekle 1994b: 171). However, other researchers have, in different contexts, employed various methods to attempt a dissection of basic shapes into smaller elementary forms. Such efforts have come from psycholinguistics, the cognitive sciences, semiotics, didactics, and, notably, linguistics (cf. an overview in Meletis 2015: 50–79). Elementary forms that are consistently assumed across these different segmentations are a (straight) line, a curve, and a dot; together, they constitute the so-called *graphetic formative lexicon* (cf. Butt & Eisenberg 1990: 36). Indeed, it is a quite trivial observation that every basic shape in the world's scripts is made up of these three components. However, the story is more sophisticated than that: for example, Primus, Fuhrhop, and other linguists have, based on a segmentation into elementary forms, attempted to show that there are inner systematics to lowercase Roman basic shapes as well as Tifinagh and Arabic basic shapes (cf. Primus 2004, 2006; Primus & Wagner 2013; Fuhrhop n.d.).

The basic shape is not only the central unit of micrographetics but arguably the central unit of graphetics in general. In the assumption of such a unit, I follow Rezec (2009, 2013), who proposed it in the context of his claim that the grapheme should be rid of its duty to serve both as a material and a linguistic unit.

46 These examples are to be taken with a grain of salt. Perceptually, and thus, descriptively, there might not be any spaces between the graphs of a word written in Arabic script or between the graphs in connected handwriting. In production, too, sequences of graphs are often written in a continuous flow and lack segmentation. Crucially, however, even writers of connected scripts are aware of the segmental basis of writing. For example, basic shapes are taught as separate units in literacy instruction, and arguably they are also stored and used as such. Segmenting a connected written word of Arabic into its respective basic shapes should thus not be a problem for users since it was conceptually and consciously composed of these segments in the first place.

47 Exceptions are single elementary forms that are simultaneously non-segmentable basic shapes, e.g. |.|, |–|, and |c|.

As a comparative concept applicable to all types of writing systems, the grapheme can reasonably be conceived of as a semiotic sign constituted by a visual unit – its signans – and a linguistic unit – its signatum (cf. Section 2.2). Both constituents are only (necessary) parts of a dyadic grapheme and can thus not be equated with the grapheme. In other words, the grapheme cannot simultaneously be the visual unit since that is only one of its constituents. It is, however, a visual sign relating a visual unit to a linguistic unit.

In his proposed optimization of a model of the German writing system, Rezec manages to divorce the functions that are commonly allocated to the grapheme: (1) being a visual unit, (2) being the smallest distinctive unit of writing, and (3) corresponding to a phoneme. He assigns the first of these functions to the so-called *basic shape* (in German originally *Grundform*). Thus, the basic shape is a material unit. It is imperative to note, however, that at the same time, it is an abstract unit. Essentially, it stores a bundle of visual features that are necessary to distinguish visually a shape from the other shapes in an inventory. As Herrick (1974: 11) already stated (long before Rezec): "The basic shape [...] is itself an abstract [...] unit; it is a group of geometrical distinctive features which a written mark must have so that a literate person will recognize it as an embodiment of a certain letter [= in the conception here, 'letter' means 'grapheme', D.M.]". What differentiates |E| from |F|, for instance, is the number of segments they consist of. By contrast, what differentiates |X| from |T| is not the number or nature of segments – in both cases it is two straight lines – but the position of these segments, the type of connection between them (crossing vs. acute angle), and, most crucially, the spatial and topological relation between them in the segmental space. To give another example, |J| and |L| are distinguished by the nature of one segment – a bow in |J| vs. a straight line in |L| –, which also influences the transition between the two segments, respectively. They also showcase a different orientation of this lower horizontal segment (leftwards in |J|, rightwards in |L|).

The linear space (see below) and with it, the space it subsumes, the segmental space, can be divided further. When four horizontal division lines are drawn, the linear space can be divided into three spaces that are vertically superimposed upon each other (cf. Althaus 1973; see Figure 5).[48] Note that this division was assumed on the basis of the Roman script and has limited applicability when it comes to other scripts (see below). The topmost of the three spaces constituted by the division lines is the *high space*, followed by the *central space* in the middle, and the *low space* at the bottom. The third of the division lines from the top – the one the basic shapes 'stand on' – is commonly also referred to as the *base line*. This division of the linear/segmental space allows a description of how exactly basic

48 Alternatively, as visualized in Figure 5 by the dotted line in the middle of the central space, the linear space can be segmented into four vertical spaces divided by five lines. This four-space schema (German *Vierlinienschema*) represents the original conception (cf. Althaus 1973). In the more modern three-space schema (cf. Domahs & Primus 2015: 133), the middle two spaces of the four-space schema together form a single space, which is referred to as *central space* (cf. Primus & Wagner 2013: 42).

shapes, at least those of Roman script, occupy the segmental space. Of great relevance are those parts of basic shapes that extend beyond the central space (which is filled by a basic shape such as |a|). Following typographic terminology, these extending parts are called *ascenders* when occupying the high space, as in |d|, and *descenders* when occupying the low space, as in |y|.

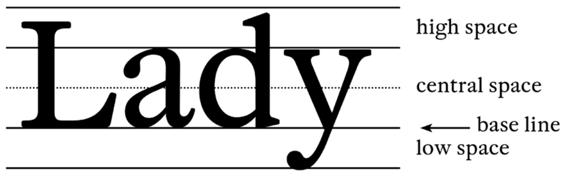



FIGURE 5. Four-space schema of the segmental/linear space in Roman script (Note: a and y are also basic shapes in Cyrillic script)

Notably, this specific spatial division of the linear space is by no means universal. Quite to the contrary, even if it applies to a number of scripts, it is fairly script-specific. For Japanese kanji and the Chinese hanzi that they are based on, for example, there exist multiple divisions of the segmental space into smaller subspaces (cf. Figure 6) depending on how the subsegmental elements of basic shapes are arranged.[49] Because of these different ways of dividing the segmental space, the segmentation in Chinese script and Japanese kanji does not extend over the entire linear space; instead, every segmental space must be subsegmented individually. A characterization of every possible segmentation of the segmental/linear spaces – complete with the identification of elementary forms and their combination to form basic shapes – is beyond the scope of this book; it is an endeavor that will need to be dealt with in detailed structural analyses of specific scripts and writing systems.

49 Another difference between scripts is that segmental spaces within a script do not have to be of equal width: in prototypical typefaces that materialize Roman script, for example, the widths of segmental spaces vary according to the sizes of basic shapes that occupy them. |i|, thus, occupies a narrower segmental space than |o|. This is mainly due to a typographic strategy referred to as *kerning*, where the horizontal distance between basic shapes is adjusted as to appear even. This is not the case in so-called *monospaced* (or *fixed-width*, *non-proportional*) *typefaces* of Roman script (such as `Courier New`) where each basic shape is assigned an equal amount of horizontal space, i.e. all segmental spaces are of equal width. The latter is also the prototypical situation in the scripts of Japanese and Chinese, where basic shapes – regardless of their complexity (including the number of strokes) – occupy segmental spaces of equal size.

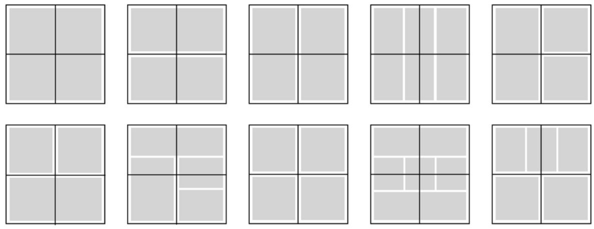


FIGURE 6. Selection of possible subdivisions of the segmental space in Japanese

So far, only units included in different scripts (Roman, Chinese, Japanese kana) were mentioned. However, writing systems make use of more kinds of visual material than just these scriptual units. Consider, for example, digits such as |2|, special characters like |$|, not to mention punctuation marks such as |;|. These are all elements of a larger group Rezec (2009: 33) categorizes as *non-letters* (German *Nichtbuchstaben*). Because graphetic research – as established above – is often located at the periphery of linguistics, and since the definition of *basic shape* is still underspecified in this respect, technically, all these mentioned units should be regarded as basic shapes. The question, now, is how one can establish that they belong to different classes. As was argued elsewhere (cf. Meletis 2015: 124f.), visually, there is no clear way of distinguishing them: by simply describing individual basic shapes such as |Z|, |2|, and |§| visually, one cannot 'see' that they belong to different classes. However, while individual basic shapes may not easily be categorized, the class they belong to can be evaluated with the help of a mixture of graphetic, graphematic, and graphotactic features, as suggested by Bredel (2011: 9). Notably, her proposal pertains specifically to the German writing system and makes no claims to universality whatsoever; unsurprisingly, it is not readily applicable to many other writing systems. However, it can still serve as a valuable example and starting point for similar future endeavors for other scripts.

TABLE 1. Classes of basic shapes evaluated with graphetic, graphematic, and graphotactic features, from Bredel (2008: 23)

| | **diacritics** | **letters** | **digits** | **special characters** | **punctuation marks** | **empty spaces** |
|---|---|---|---|---|---|---|
| *identifiable without context* | + | + | + | + | + | – |
| *recodable* | – | + | + | + | – | – |
| *combinable* | – | + | + | – | – | – |
| *paired* | – | + | – | – | – | – |
| *additive* | + | – | – | – | – | – |

As evident from Table 1, Bredel proposes five features for the distinction of different classes of segmental graphetic material occurring in the writing system of German: letters, digits, special characters, punctuation marks, and empty spaces (and, in an earlier work, diacritics). The features themselves are (1) *context-free identification*, (2) *recodability*, (3) *combinability*, (4) *paired variants*, and (5) *additivity*: (1) context-free identification is a graphetic feature as it is determined visually, (2) recodability is a graphematic feature based on the linguistic units that basic shapes are related to, (3) combinability is a graphotactic feature, and (4) paired variants is, depending on the view taken, either graphematic, as there is often no visual similarity between uppercase and lowercase basic shapes (e.g. |A| and |a|) and they are just paired according to the linguistic units they correspond with (cf. Section 2.3.3), or conventional, when the pairing of corresponding uppercase and lowercase basic shapes is treated as a convention. (5) Additivity is also graphetic in nature; it characterizes diacritics, i.e. smaller and dependent segments of basic shapes that attach to basic shapes within a single segmental space.

The only class of graphetic material that is not identifiable without context is empty spaces, which are made visible only by non-empty graphetic material around them. The only feature of punctuation is that it is identifiable without context.[50] It does not display any of the other features: it is not verbally recodable, which means it is usually not ‘verbalized’ or ‘read’ the way the grapheme <b> can be read as [b] or the special character <%> can be read as [pɜɹ’sɛnt].[51] Furthermore, punctuation marks cannot combine with each other: the ellipsis <…> is interpreted as one mark, and while there are exceptions such as <?!>, punctuation marks generally do not combine freely with one another to productively form new units the same way digits or letters do, e.g. <27> or <twenty-seven>. Lastly, punctuation marks (as well as digits, special characters, and empty spaces) are, unlike letters, not available in two different variants: letters are, at least in Roman script, which Bredel’s work focuses on, available in lowercase and uppercase variants,[52] whereas punctuation marks are not.[53] Again, it is crucial to note that these

---

**50** Note that it might not always be the case that punctuation is identifiable without context: take the comma, which (most often) shares its shape with the apostrophe. Without a context, it cannot be evaluated in which vertical subspace of the linear space the shape is located and whether it functions as a comma or an apostrophe. However, this only concerns the individual identity of a basic shape (and its graphematic function), as in any case, the basic shape will be identified as belonging to the class of punctuation (since both comma and apostrophe are punctuation marks).

**51** This feature determines that the slash </> is no punctuation mark – at least not in German – on the basis that it can be verbalized (cf. Bredel 2009: 119). An example of this is one of the variants of writing genderwise correctly in German (cf. Section 8.3), e.g. in <Student/innen>, which is to be read as <Studenten *und* Studentinnen> ‘male students *and* female students’. Here, the slash is verbalized as the conjunction *und* ‘and’.

**52** Most other scripts and, in turn, writing systems do not have this distinction between uppercase and lowercase basic shapes. In these writing systems, thus, there might be no feature distinguishing the class of digits from the class of ‘letters’.

**53** Evidently, some punctuation marks such as parentheses <( )> or quotation marks <“”> are also in a way ‘paired’. These pairs, however, always necessarily occur together, unlike <A> vs. <a>, which do not and are thus more reasonably conceived of as variants (allographs) of a more abstract unit, a grapheme (cf. Section 2.3.3).

features have not been tested for the graphetic module of writing systems other than German. Nevertheless, it is expected that they hold for all alphabetic writing systems that have a case distinction.

Arguably, terms such as *letter* or *character* are not appropriate when used as designations for language-specific graphemes. *Letter* is currently being used this way for the basic units of many writing systems, not only for alphabets but also for abjads (cf. Section 2.7), while *character* is strongly associated with the units of the Chinese writing system as well as systems that have developed from it. This use, however, is misleading, as it obscures the relevant features that graphemes of different writing systems share. Instead, these terms should be reinterpreted as graphetic terms that by convention designate classes of basic shapes that are part of certain scripts. Note that for some scripts, there may not even be terms comparable to *letter* or *character*, which is when the general *basic shape* can prove useful.

Due to the lack of a universal heuristics that allows distinguishing different classes of basic shapes across writing systems the way Bredel's (2011) above-mentioned criteria allow for German and related systems, I argue that what is vital for an analysis across systems is knowledge about the different classes. This knowledge works top-down and allows categorizations (cf. Meletis 2015: 124f.). In other words: when readers and writers are proficient in a writing system, they know what class a given basic shape belongs to on the basis of knowing which linguistic unit it relates to and in which contexts it is used.[54] As there is usually no overwhelming systematic visual coherence within the shapes of the classes to tell the classes (i.e. scriptural basic shapes vs. punctuation marks vs. digits, …) apart,[55] this knowledge is predominantly graphematic. To summarize, *basic shape* is generally an underspecified term that designates, at an abstract level, all visual units used in a writing system. However, since this book is primarily concerned with basic shapes of the type *letter* or *character*, i.e. basic shapes that are the visual parts of a writing system's default graphemes, the unmarked use of the term here is restrictive and means only them. A possible, though flawed specific term could be *scriptual basic shape*,[56] insinuating that these basic shapes are part of a script inven-

54 It is also this knowledge that tells the reader whether an element even belongs to one of the classes of basic shapes or not. In other words: whether something even *is* a basic shape of a writing system, i.e. writing, or rather a drawing, a scribble, or a sign (such as an emoji, for instance) that is not part of the writing system *proper*.

55 A part of it might be graphetic knowledge, too, as scripts are visual systems with specific characteristics: 'letters' of the Roman script, for example, have their coda prototypically on the right side – |b| or |D| – while digits prototypically lean to the left: |3| or |9|. Each inventory that has existed for a longer period and that thus has had time to develop usually exhibits a certain degree of systematicity (cf. Watt 1983a). This systematicity is at the core of the so-called systematic fit, cf. Chapter 5.

56 This term highlights the fact that these basic shapes are elements of a *script*, a (most often) closed inventory. It is more general than the script-specific designations *letter*, *character*, etc. It also avoids mixing the linguistic and the material levels. However, it is an undeniable fact that these scriptual basic shapes are the ones that are used to

tory, whereas digits, punctuation marks, and special characters are not. This is also reflected by the fact that these latter classes – especially punctuation marks – are used across many writing systems regardless of the scripts (e.g. Roman, Cyrillic, Chinese) employed by these systems.

While the assignment of basic shapes to classes is not a visual, and thus, not a graphetic matter, the differentiation between different individual basic shapes is. Thus, |F| and |E| are different basic shapes mainly for the reason that they differ visually and not because they relate to different linguistic units and therefore, are parts of distinct graphemes (but see below for |T| and |Γ|). Vice versa, visually distinct and thus separate basic shapes such as |ς| vs. |σ| can be assigned to the same grapheme – in this case, the grapheme <ς/σ> that in Greek refers to the phoneme /s/. |ς| and |σ| are allographs of a grapheme, but they are still distinct basic shapes (and not just different graphs, see below) (cf. Section 2.3.2). Thus, it is the abstract visual information stored in a basic shape – as a visual common denominator – that is distinctive.

The abstractness of the distinctive visual information leaves a lot of leeway for graphetic variation: the 'visual skeleton' (cf. Cox et al. 1982) that constitutes the basic shape can be materialized in countless different ways (cf. Figure 7). This explains how, for example, different people's handwriting or different typefaces have specific visual characters and are perceived as 'different' but their units can still be identified and assigned to respective basic shapes. Indeed, the human visual and cognitive systems allow us to recognize graphs that look different and categorically assign them to basic shapes as long as they are located within the respective graphetic solution spaces (see below), i.e. do not resemble a different basic shape that is possibly part of a different graphematic relation or a shape that is not a basic shape at all. This leads to a necessary definition of the unit at the lowest level of writing, the *graph*.

A grapheme needs to have as its visual component a basic shape. In some cases, it has more than one possible visual constituent, as is the case for the above-mentioned Greek <ς/σ> that is visually manifested by |σ| and |ς| depending on its position in a word. A basic shape, in turn, at the level of the concrete realization of writing, needs to be materialized by graphs. Thus, while the basic shape is an abstract and to a degree theoretical unit, the *graph*[57] is its concrete material instantiation (cf. Adam 2013). Every graph is a unique physical event. Every graph is also always an allograph, as it is only one of countless possible realizations of

embody the default graphemes of a writing system, while other basic shapes – digits, special characters, and punctuation marks, whose designations are also flawed since they are derived from their functions – are peripheral. Note, however, that the term *scriptual basic shape* is suboptimal insofar as these other classes of basic shapes, e.g. the set of digits, could also be interpreted as 'scripts', i.e. closed inventories of visual shapes. However, these sets or 'scripts' in the broader sense are not constitutive of writing systems.

**57** Sometimes, the alternative designation *glyph* is used, a term borrowed from typography (cf. Neef 2015: 711). *Graph* and *glyph* can be considered synonyms.

the same basic shape (cf. Section 2.3.1). Because it is the concrete visual level of writing, the level of graphs is the scope of all visual variation in writing. As Ludwig (2007: 382) correctly observes, there are (seemingly) no limits to this kind of variation. Indeed, as long as the abstract visual features of a basic shape – most importantly the number of segments, the arrangement of segments in space, and the topological configuration of segments with respect to each other – are kept relatively constant, everything else can vary.

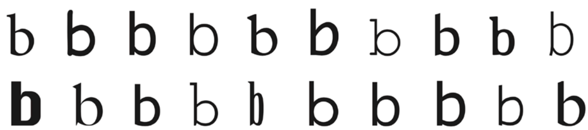


Figure 7. Basic shape |b| materialized by twenty different graphs in different typefaces

To illustrate this, consider Figure 8, in which the basic shape |A| is materialized as a prototypical graph (left) and two other versions in which visual features have been distorted, respectively. While the middle graph, in which the relative length of the segments has been altered, is still recognizable as a realization of |A|, the rightmost version, in which the topological configuration of elements was changed, is, at least if presented in isolation, possibly too distorted to recognize as |A|.

Figure 8. Geometrical vs. topological distortion of a prototypical basic shape, from Meletis (2015: 164)

By analogy with Neef's (2005, 2015) graphematic solution space, in which different spellings for phonological strings (and, in an extension of the concept to morphographic writing systems, morphemes) are located (see the introduction to Chapter 2), a *graphetic solution space* is proposed here (cf. Blesser et al. 1974 for a similar concept, the 'character space'). In short, the licensed variation manifested by graphs that are assigned to the same basic shape is located within the boundaries of a basic shape's graphetic solution space. Thus, the graphetic solution space is the descriptive counterpart of what makes possible visual categorical perception and, for that matter, optical character recognition (OCR). The variants within this space may differ only very subtly. Even if a person (with sufficient graphetic awareness) perceives the individuality of graphs within the graphetic solution space, the differences between them are still non-distinctive at the level of basic shapes and, in turn, graphemes.

An issue that is central to the graphetic solution space is the investigation of these differences between graphs, and an evaluation of how much given graphs differ. As established, the graphetic solution space for a given basic shape includes all the graphs – as concrete materializations of the abstract basic shape – that are visually categorized as being members of said basic shape. This graphetic solution space is crucially dependent on both the script (as the visual system that the basic shape is a part of) as well as the entire writing system (as a linguistic semiotic system) that employs this script. The boundaries of the graphetic solution space, thus, are determined graphematically, i.e. not visually but linguistically. In Greek script as used for the Modern Greek writing system, for example, there is a categorical distinction between the basic shapes |Τ| and |Γ| since they are part of distinct graphemes, i.e. they are used to correspond with different linguistic units (the phonemes /t/ and /ɣ/). Thus, the graphetic solution space for them will not be as large and as 'forgiving' as the solution space might be for |T| in Roman script as used by many alphabets (cf. Figure 9 for the graphetic solution space of |Τ| in Greek). Notably, in the latter, the basic shape |Γ| does not exist, and thus, is not graphematically associated with any linguistic unit. Therefore, in Roman script, the graphetic solution space for |T| is larger because there is no danger of mistaking it for |Γ|. From this also follows that it is impossible to assume "distinctive features" of a script by purely graphetic means, as what is visually distinctive relies on what is graphematically distinctive. Furthermore, a given feature might be distinctive in some instances in a script used by a specific writing system and non-distinctive in others. Accordingly, neither script-internally nor universally, i.e. across scripts, is there a way to determine an inventory of distinctive features of basic shapes. Distinctions are only meaningful if any two basic shapes of a script and their relationship with each other are considered, as illustrated by the example of |Τ| and |Γ| in Modern Greek. The length and/or position of the upper stroke is distinctive in this case, but these same features might not be distinctive in any other two basic shapes of the same script.

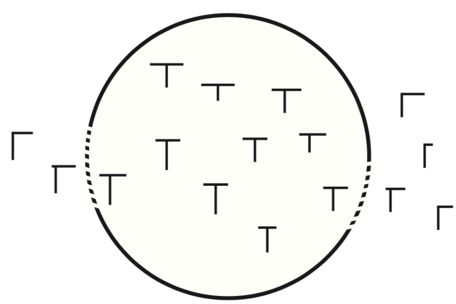


Figure 9. Graphetic solution space for |Τ| in Greek

Graphematics imposes boundaries onto the graphetic solution space(s) in a top-down manner only in cases in which visually, graphs are getting too similar to a distinct basic shape and this similarity could, at the graphematic level, lead to a wrong categorization. Inversely, there also exist visual distinctions in the graphetic solution space that do not correspond to graphematic distinctions and are, thus, independent of graphematics: take |g| vs. |ɡ| in writing systems that use Roman script. They are characterized by a visual distinction between them. Notably, this distinction may be hard to perceive for users given that they know that these two shapes both ‘signify’ the same, which, in graphematic terms, means they relate to the same linguistic unit. When asked, users (in this case of English) often do not even know that <g> has two variants (cf. Wong et al. 2018). Crucially, here, the visual distinction does not correspond with a graphematic distinction (cf. Section 7.1.3). However, because of the visual dissimilarity and possibly also because of the top-down conventional knowledge that both are existing variants of one abstract unit, |g| vs. |ɡ| are distinct basic shapes.

Note that what has been established so far concerns the perception and recognition of individual and isolated basic shapes. In the context of a sequence of basic shapes that manifests a graphematic word, such as <ΛCCESS>, even major distortions such as the omission of segments as in |Λ| for |A| might be ‘forgiven’ perceptually because of disambiguating information offered by the context. This corresponds with how perception is modeled in the influential *Interactive Activation Model* (cf. McClelland & Rumelhart 1981; Section 7.1.3). Larger contexts will be addressed in the next section.

When basic shapes (or better: concrete graphs) are produced next to each other, each of them occupies its own segmental space[58] and they are written in a row, whether horizontally (from left to right or right to left) or vertically (from top to bottom or, very seldom, bottom to top). These ‘rows’ lead to the first polysegmental level of writing: the linear space that is at the center of mesographetics.

### 1.2.2 Mesographetics: one-dimensional graphetic sequence, line

Two graphetic units occupy the linear space and are studied by mesographetics. They are distinguished from each other in that they are constituted by different empty spaces. Also, while the first of them fills only part of the linear space, the second one fills all of the available linear space: the first is the *one-dimensional graphetic sequence*, the second the *line*. Note that the line is, technically, just a special case of a one-dimensional graphetic sequence in which all or at least most of the linear space is occupied by graphetic material. While the one-dimensional graphetic sequence is, in most cases, a graphetic unit only secondarily since it is frequently determined by the graphematic level, specifically the type of linguistic unit (morpheme, lexeme, phrase, etc.) that is being visualized, the line is, with only few exceptions, a purely graphetic unit.

58 However, sometimes, they are ‘shrunk’ in size and become elementary forms themselves, occupying segmental spaces only together with other ‘shrunk’ basic shapes, as in Chinese.

The one-dimensional graphetic sequence only exists in writing systems in which there are spaces either between words[59] or syntactic units. The latter is the case in Thai, for example, where, due to the lack of empty spaces between words, one-dimensional graphetic sequences visualize and/or are constituted by syntactic units. In all writing systems using Roman script or a modified version of it, the one-dimensional graphetic sequence commonly manifests words. However, it is crucial to note that a correspondence with the lexical or morphosyntactic definitions of 'word' or 'sentence' – whatever these may be in a given context – is not required for the definition of the one-dimensional graphetic sequence, which is an entirely independent graphetic unit: everything that stands between two empty spaces of this order and consists of at least two basic shapes (i.e. occupies at least two segmental spaces) regardless of their class qualifies as a one-dimensional graphetic sequence. This means that in alphabets, basic shapes such as, for example, punctuation marks – whether at the word-level such as <'> or at the sentence-level such as <!> – that are attached enclitically to other basic shapes are parts of one-dimensional graphetic sequences. Thus, at the end of the preceding sentence, |sequences.| (highlighted in grey) is a one-dimensional graphetic sequence, and in the preceding phrase, |sequence,| is a one-dimensional graphetic sequence. Visually, in these cases, the spaces between basic shapes and punctuation marks are not larger than between the basic shapes themselves. Note that these units are not graphematic words as defined in German grapholinguistics, as (sentence-final) periods and commas are, unlike punctuation marks at the word level (such as the apostrophe and the hyphen), not regarded as parts of graphematic words (cf. Fuhrhop 2008: 217; Evertz 2016: 391; Section 2.5). One-dimensional graphetic sequences and graphematic words are also incongruous in cases in which graphematic words consist of only one grapheme, such as the article <a> in English or the conjunction <y> 'and' in Spanish. These units occupy only one segmental space and are, thus, not one-dimensional graphetic sequences. Yet, if analyzed graphematically, they are graphematic words.

The second unit that extends in the linear space is the line. It meets the empty space criterion because it is visualized by the line break. Lines are constituted by the fact that on many writing surfaces, scribes eventually reach a phys-

59 Previously, in a script-graphetic analysis of the visual material used in the German writing system, this unit was referred to as 'graphic word' (cf. Meletis 2015: 130–132). On the one hand, this is problematic because it mixes the graphematic and the graphetic levels of description. Even if the one-dimensional graphetic sequence corresponds largely with 'words' (however one defines that linguistic unit or category), this correspondence cannot be constitutive of graphetic units. On the other hand, a term such as 'graphic word' is inherently specific. It cannot be used for writing systems in which one-dimensional graphetic sequences do not correspond with words but with other linguistic units. Instead of assuming different units such as 'graph(et)ic word' and 'graphetic sentence' based on their correlations – which is done in a graphematic analysis – we are only concerned with the units' visual features at this point. The common visual denominator of these units regardless of any linguistic correspondence is that basic shapes that occupy segmental spaces are produced in a non-spaced sequence. This sequence occupies a linear space.

ical boundary, often the edge. Reaching this edge commands that one continues writing a little bit below (in horizontal top-down writing systems) or to the left of (in vertical right-left writing systems) the last line. Unlike the one-dimensional graphetic sequence, the line is commonly not functionalized as a graphematic unit. It can relate to linguistic units in the sense of loose semantic units when considering verses in poetry, which, however, commonly also only occupy a part of rather than the entire linear space. Notably, line breaks can be intentional, for example in typography, where they are sometimes aesthetically motivated. In this case, line breaks are conscious choices made by the writer/designer[60] that have nothing to do with the physical boundaries of the surface. Verses or aesthetically motivated lines, of course, are not the default types of lines.

### 1.2.3 Macrographetics: two-dimensional graphetic sequence, page/layout

When linear spaces are concatenated (either horizontally or vertically, but necessarily in the other dimension than the respective concatenation of segmental spaces to form linear spaces), the resulting clusters of lines constitute two-dimensional graphetic sequences. The space that they occupy is the areal space, and the empty space that makes them visible is located between areal spaces. Two-dimensional graphetic sequences can be functionalized differently, examples being paragraphs or columns. These, notably, are not graphetic categories since, for the assumption of graphetic categories, visual criteria must suffice, and it is debatable how paragraphs and columns can be distinguished by visual means.[61]

The next 'unit' is constituted by the arrangement of two-dimensional graphetic sequences as well as other visual material such as photos and figures on the entirety of the writing surface. If the surface is a page on which several paragraphs, a few footnotes (i.e. paragraphs at the bottom of the page in smaller print), and maybe a figure or a table are printed, as is the case on many pages in this book, these are all elements that are spatially arranged in the so-called holistic space. This arrangement of elements in the holistic space is commonly referred to as *lay-*

60 Ludwig (2007: 377) notes that in the past, the tasks of writing a text and designing it were undertaken by different people with distinct professions. Even though many people nowadays work with word processing programs and not only write but also format/design their own written products, these tasks are in many contexts still separated. Authors who hand in manuscripts of their books to publishers, for example, often do not participate in the formatting process (at least not the final, professional formatting process). These different tasks and the associated professions also reflect the underlying distinction between graphetics and graphematics.

61 A possible answer is that in columns, in horizontally written writing systems, lines commonly do not fill all of the linear space on the 'page' (see below) but only part of it, while in paragraphs, they prototypically fill most or all of the linear space. Also, columns, as two-dimensional spaces, are typically concatenated next to each other on the horizontal axis, and not below each other like paragraphs.

*out*. It is important to note that holistic spaces are not just pages or double pages, the latter of which are often perceived simultaneously when reading a book, but any writing/reading surface that can be perceived at once. For example, this can also be the section of a website that is currently displayed on a screen or a wall on which PowerPoint slides are being projected. When scrolling up or down or pressing a key, the holistic space in these examples changes. In this sense, holistic spaces are dynamic and determined by what is perceived by users as a "whole" space. The printed page is only the prototypical analog version of the holistic space.

A phenomenon relevant at the macrographetic level is so-called *typographic dispositifs* (cf. Wehde 2000), which could be more generally termed graphetic dispositifs. If the arrangement of elements on a page allows the perceiver to immediately identify the genre of a text, it functions as a graphetic dispositif. If a text is designed rather prototypically, i.e. its elements are arranged in a way typical of a genre, it should be easily recognizable as, for example, a recipe or the front page of a newspaper, even if the content is replaced by X's and the text thus cannot give hints (cf. Figure 10). What counts and works as a graphetic dispositif is, of course, utterly culture-specific and in general determined by an abundance of factors (such as epoch, region, familiarity of a genre, graphetic knowledge on behalf of users, etc.).


FIGURE 10. Graphetic dispositifs of a recipe (left/middle) and a newspaper front page (right)

An insightful macrographetic study that proposes a method of distinguishing running texts from lists by visual means is Reißig (2015). In his study, Reißig aims to show that graphetics and syntax are connected. To accomplish this, he operationalizes several concepts that were originally devised in the field of typography. In what he terms the cartography of the medial (under)ground, he vertically divides the page into three equally wide list spaces: left – middle – right. Now, when the items of a list are not visually marked by bullet points or numbers, what is crucial for users to distinguish between lists and running texts is how much of the linear space is filled by list items: just the left list space, or does the text run beyond that? To conceptualize this visual difference and separate the list mode from the text mode, Reißig (2015: 33–35) proposes the feature [±CONTINUOUS]. Lines that occupy not only the left but also the middle and right list spaces are [+CONTINUOUS], while lines that occupy only the left or the left and (limited parts of) the middle space are [−CONTINUOUS]. These gradual feature values are visually salient, and what is perceived by a reader as a list or as running text is a matter of graphetic dispositifs. In a follow-up study, Reißig & Bernasconi (2015: 235) tested the perception of these graphetic dispositifs empirically and arrived at the conclusion that with the decreasing length of lines, i.e. the decreasing occupation of the linear space by graphetic material, readers decide in favor of the list mode. Not only does Reißig's contribution enrich grapholinguistic theory, but it also constitutes important evidence for the claim that writing is a system in its own right, as the list mode proves that in writing, due to its spatial nature, there are modes of organization that have no equivalent in speech.

### 1.2.4 Paragraphetics

Micrographetics, mesographetics, and macrographetics treat graphetic phenomena that are perceived two-dimensionally. However, as established above, one of the central features of writing that distinguishes it from speech is that it requires a surface and tools (cf. Dürscheid 2016: 31). These aspects should also be considered and studied in graphetics. Since a product of writing reveals the material it was made of and often exhibits traces of the tools and methods it was made with, Stöckl (2004: 37–39) proposes an additional level of analysis that includes the third dimension: paratypography, which, for the sake of generality, can be re-termed paragraphetics. Stöckl chooses the prefix *para-* for this level because the aspects of writing treated by paragraphetics affect the entire process of producing and perceiving a product of writing.

The physical features of the writing/reading surface greatly influence processes of writing and reading. Possible properties studied here include the initial choice of paper or other materials as well as their color/brightness, transparency/opacity, surface (matt vs. glossy), grey-scale value, and haptic phenomena such as their thickness, density, grammage/weight (cf. Spitzmüller 2016: 101f.; Willberg & Forssmann 2010: 71; König 2004: 97f.) but also external factors such as the incidence of light when writing or reading, to name only a few. As König (2004: 73f., my translation) puts it, "optimal typographic readability emerges from the best possible interplay of individual typographic factors with simultaneous consideration of the reception situation and the individual reader".[62] If, for example, in a given reading situation, the transparency of the paper colludes unfavorably with the light, the reading process might be hindered to some degree. Ziefle (2002: 50–61) extends the study of these factors to reading on computer screens and shows that contrast/lighting, resolution, and flickering are relevant categories and that generally, reading on paper offers better conditions than reading on screens.

62 „Die optimale typographische Lesbarkeit ergibt sich aus dem bestmöglichen Zusammenspiel einzelner typographischer Faktoren unter Berücksichtigung der Rezeptionssituation und des individuellen Lesers."

While paragraphetic considerations are far from being linguistic, they are of the utmost relevance when it comes to studying reading and writing.

## 1.3 Towards a typology of scripts

One of the most prominent and well-studied subfields of grapholinguistics is writing system typology. Despite its broad name, it deals almost exclusively with the question of which linguistic units the graphemes in different writing systems relate to. This is evident from the designations of assumed types such as *morphography* or *syllabary*. This question is a graphematic matter, which is why writing system typology – which amounts to 'graphematic typology' – will be discussed in the chapter covering graphematics (cf. Section 2.7). Crucially, now, as they focus on graphematic relations, different established typologies disregard visual, i.e. graphetic considerations completely. Indeed, writing systems that are typologically related do not necessarily use visually similar scripts, and inversely, just because the scripts used by two writing systems resemble each other does not mean that the writing systems have anything in common functionally. In line with arguing that the visual aspects of writing should not be neglected, however, this section raises the question of how scripts, i.e. visual inventories, could be categorized into types based solely on graphetic features.

At the outset, it must be said that there is no such thing as a 'script typology' yet. One could falsely assume the opposite due to the different uses of the term *script*.[63] However, if it is read the way it is in this book, i.e. as an inventory of visual basic shapes that can be used for the materialization of a (in fact *any*) writing system – examples being Roman script, Chinese script, and the kana inventories of Japanese –, then no actual typologies exist. A laudable, though effectively misguided effort at a 'taxonomy of alphabets and scripts' that is worth mentioning in this context was made by Earl M. Herrick (1974: 5), who stated over forty years ago:

> Linguists, typographers and others who work with written language do not presently have any adequate system of classification for describing similarities and differences among […] scripts which are used to write languages.

With his taxonomy, Herrick attempted to change this situation. Although he mentions both linguists and typographers in his introduction, the latter of which are predominantly interested in the visual aspects of writing, he strays far from the definition of *script* adhered to here and arrives at an unfortunate mixing of the material and linguistic levels. This becomes particularly obvious when he lists three criteria of 'alphabetical similarity' for arriving at a taxonomy of scripts:

---

63 Gnanadesikan (2017), whose findings will be central in the section on writing system typology (cf. Section 2.7), calls types of writing systems *scripts*. For example, she calls the phonographic type *phonemic script*. This does not correspond with the definition of *script* as adhered to in this book. What I call *script* is, for Gnanadesikan (2017: 15), a *signary*.

1. Similarity in the basic shapes of letters.
2. Similarity in the correspondences between letters and phonemes.
3. Similarity in the alphabetical orders of letters. (Herrick 1974: 12)

In a 'pure' script typology, i.e. one based solely on visual features, only the first criterion should matter. The second would already be part of a typology of writing systems in which, conversely, as mentioned above, the first criterion is ignored.[64] Note the interesting terminology Herrick employs here: for him, *basic shapes* are exactly what they are in the present approach, while he uses *letter* as a functional term, probably as a synonym of the analogical reading of *grapheme* (cf. Section 2.2). That he follows the analogical rather than the referential view becomes apparent in his second criterion, with which he basically invokes grapheme (or letter)-phoneme correspondences. His third criterion is interesting because it is relevant neither for a typology of scripts nor for a typology of writing systems. Examining the order of the basic shapes in an inventory is a historical matter, and ultimately a matter of convention. Consequently, the order in which elements are grouped in a script or writing system[65] does not have an impact on the fundamental criteria that allow assuming typologies of scripts on the one hand and of 'writing systems' (i.e. graphematic modules) on the other.

Let me illustrate an example of how Herrick's taxonomy works in practice. The taxonomy's hierarchical levels are the following: *alphabets*[66] < *scripts* < *genera of scripts* < *families of scripts*. So, first, Herrick groups several 'alphabets' together, for example, Czech, Dinka, English, Hawaiian, Icelandic, Kazakh, Malay, Navajo, Serbo-Croatian, Spanish, Swedish, and Turkish, and assigns them to a 'script' that he calls 'Neoroman'. The assignment of these alphabets to one script

64 Interestingly, around the same time, Herrick (1975a: 537), in a different contribution, wrote: "A linguist should be able to analyze a script, and the contrastive features which distinguish its marks from one another, without having to deal at the same time with the linguistic features peculiar to any one language". Ironically, in his 1974 taxonomy, he had failed to do exactly that.

65 It is not obvious which of the two is ordered: is the conventional order inherent to scripts or writing systems? I would argue that it is the first, as the ordering of units is not dependent on the specific writing system. This is obvious as it stays consistent across various writing systems that use the same script: for example, 'letters' of different alphabetic systems using Roman script have more or less the same order, with the exception of units that have been added – for example |ü| in German or Turkish, |å| in Scandinavian versions of the script, or basic shapes modified by hačeks such as |č| in versions used by Slavic writing systems. These additional units must be integrated into the order of the initial, original set.

66 Herrick's mixing of the material and functional levels is nowhere clearer than in his use of *alphabet*. The 'smallest' individual inventories that Herrick considers – Czech, English, German, etc. – are in fact alphabetic writing systems. They do not differ in their use of script, which is exactly why Herrick groups them together as 'scripts', specifically under 'Romanoid'. However, they are identified as different Romanoid scripts in Herrick's conception since they represent different writing systems that serve different languages. Following Weingarten's (2011) analysis of writing systems as combinations of scripts and languages, the base level of Herrick's *script* typology is determined by different languages instead of different scripts.

is based on the fact that “letters must have almost the same alphabetical orders and almost the same pronunciations, and they must be embodied almost entirely by characters with identical basic shapes” (Herrick 1974: 16). The same criteria are applied for the assignment of scripts to genera and genera to families. However, the degree of similarity decreases: for two alphabets to be assigned to one script, Herrick uses the attribute “almost the same”, whereas, for two scripts to be assigned to a genus, this changes to “similar”. The ‘Neoroman’ script is grouped together with the scripts ‘Paleoroman’ (consisting of the alphabet Latin), ‘Fraktur’ (consisting of German), and ‘Irish’ (consisting of Gaelic), and together they form the genus ‘Romanoid’. Finally, at the hierarchically highest level, Herrick groups together the three genera ‘Romanoid’, ‘Cyrilloid’, and ‘Hellenoid’ and calls the resulting family ‘Hellenic’. What this classification implies is that Herrick’s unique blend of criteria leads to a taxonomy that highlights genetic affiliation. It is not so much a typology of scripts, then, in the traditional linguistic sense of ‘typology’, but more of a genealogy of scripts and writing systems.

In sum, because of his flawed choice of criteria and his mixing of material, functional, and conventional aspects, Herrick’s proposal cannot be regarded as a typology of scripts. In his suggestions for future research, he claims that his taxonomy is only preliminary and that it “will require much work by many hands to become a useful taxonomy of writing systems” (Herrick 1974: 29). In finally using the term *writing system* in his closing remarks, Herrick reveals that his distinction between *script* and *writing system* is fuzzy and that he does not appear to be sure of what exactly the results of his endeavor represent.

In an elaborate volume on the diversity of language and writing titled *Der Turmbau zu Babel* (transl. *The Tower of Babel*), an overview of the ‘writing systems of the world’ (Seipel 2003: 10–11) is provided. Similar to what Herrick arrived at, it represents a genealogy (cf. Figure 11) that highlights the historical connections between scripts. Such a genealogical tree may be called upon, for example, when the striking visual similarity of two scripts must be explained. However, as mentioned above, as for languages, for scripts, too, genealogy does not equal typology. What even *is* a typology of scripts, then, and what are the relevant criteria that help assemble it?

When scripts are defined as visual inventories devoid of linguistic information, this means categories that are purely visual must be provided for a script typology. There are multiple candidates for a visual base criterion, all of which lead to different typologies. Before these can be assessed, however, a question that needs to be asked is what use a script typology would have and how grapholinguistics (and a theory of writing) could benefit from it.

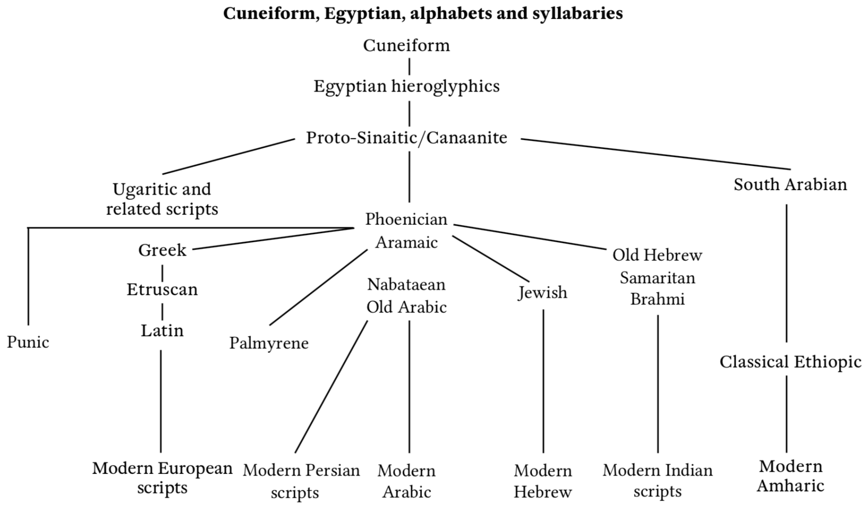


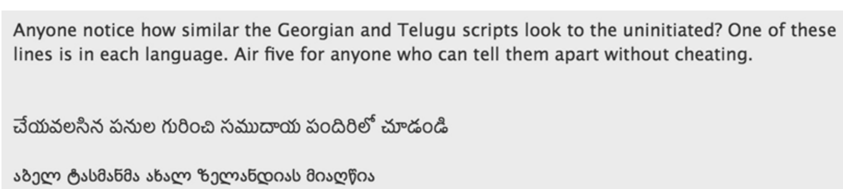

FIGURE 11. Genetic classification of scripts, adapted and translated from Seipel (2003: 10–11)

The kind of script typology that is proposed here is *descriptive* as it is based on descriptive visual criteria (as opposed to usage-based criteria). As such, it falls into the realm of script-graphetics as presented above. This corresponds with the nature of existing typologies of writing systems, which are equally descriptive in nature. If description is the starting point, a different perspective can be adopted in the next step, and the reality of descriptive categories for the processes involved in the production and perception of scripts can be studied. In this vein, a descriptive typology of scripts is a deductive predecessor to (and, arguably, a prerequisite of) a functional theory of scripts. It will be interesting to test whether the categories that were proposed descriptively actually have a bearing on the physiological and cognitive processing of scripts, which will be assessed in Section 7.1.

Two different types of visual categories must be distinguished according to the level at which they are located: at the *micro-level*, features of individual basic shapes are in focus. What is analyzed, thus, is a basic shape within its segmental space. The features relevant at this level are *individual features*. Arguably more crucial for the typology of scripts, however, is the *macro-level* at which *relational features* of a script's basic shapes are examined. For instance, how distinctive a basic shape is can only be assessed when considering its relation to the other basic shapes of a script. In sum, the relational features of scripts lead to an overall value of systematicity that, among other things, allows an evaluation of how 'natural' given basic shapes are within a script or what potential new basic shapes that are being added to a script could look like (cf. Watt 1983a; cf. Section 5.1).

Many of the potential visual categories for a script typology are informed by the spatial arrangement of basic shapes. When only a single segmental space and the individual features of a given basic shape are investigated, several questions arise, some of them of quantitative, others of qualitative nature. Possible quantitative questions are: how many elementary forms (straight lines, curves, dots) does a basic shape consist of? How much of the segmental space is filled by visual material (in contrast to empty space)? Qualitative questions concentrate on the organization of the visual material: how can a basic shape be segmented, i.e. *which* are its elementary forms? How are these arranged within the segmental space, and is the segmental space internally structured in a complex way, too, as exemplified by the four-line-schema in Roman script or the possible subspaces that subsegmental components can occupy in Chinese hanzi and Japanese kanji (see above)? Are the elementary forms of the basic shape hierarchically organized in a way similar to how Primus (2004, 2006) proposed for Roman script? In this vein, is there a hasta/head and a coda? Although the following is primarily a relational feature relevant for the analysis of an entire script, it can also be asked when the spatial arrangement of an individual basic shape is assessed: does a basic shape have a direction? Is a basic shape intrinsically symmetrical, like |A| or |o|? What is the topology of the elementary forms in relation to one another, i.e. how and where in the segmental space are they connected? How can, for example, the difference between |T| and |L|, basic shapes that consist of two straight lines, respectively, be described and formalized?

As implied, even before they contribute to a hypothetical typology of scripts, these questions can be used to describe individual basic shapes and compare basic shapes with each other. Questions that regard the whole script and lead to a discovery of relational features are, among others: is there extrinsic symmetry, i.e. are there basic shapes that are treated as distinct in the script but differ only in orientation, such as |p| and |q|? What is the overall 'character' of a script: is it more angular or round? Consider, for example, the overall visual character of the scripts in the two figures below. While in the scripts of Georgian and Telugu (cf. Figure 12), basic shapes are overall curved, as noted by a user in an online forum, the Chinese hanzi as well as the basic shapes they have influenced – the Japanese kana inventories[67] and Korean Hangul – are predominantly angular (cf. Figure 13).


Anyone notice how similar the Georgian and Telugu scripts look to the uninitiated? One of these lines is in each language. Air five for anyone who can tell them apart without cheating.

చేయవలసిన పనుల గురించి సముదాయ పందిరిలో చూడండి

აბელ ტასმანმა ახალ ზელანდიას მიაღწია

Figure 12. Roundness of the Telugu (above) and Georgian (below) scripts, from https://forum.unilang.org/viewtopic.php?t=35659 (July 25th, 2020)

67 The kana inventory pictured in Figure 13 is *hiragana*. In Japanese, it was previously also called *onnade* 'women's script' because it was predominantly used by women when men were using Chinese characters (cf. Coulmas 1996a: 207). Another reason it was called 'women's script' concerns the fact that it is visually more curved and thus 'smoother' in nature than the more angular *katakana*, which are not pictured (cf. Dürscheid 2016: 87; Stalph 1996: 1420).

| | | |
|---|---|---|
| 문자는 언어를 기록하기 위한 상징 체계이다 | 文字是一种语言的标示系统 | 文字は言語を表す符号である |

Figure 13. Angularity of the Korean Hangul (left), the Chinese (middle) and the Japanese kanji and katakana scripts (right), adapted from https://storage.googleapis.com/spec-host-backup/mio-design%2Fassets%2F1Fv3Pph5qFaeOLFhbXTPDTo0OYAaalvbA%2Flanguage-2.png (July 25th, 2020)

The questions listed above concern the micro-level as well as the macro-level and can be bundled systematically as done in Table 2. After determining the individual feature values for each basic shape in a script, the result can be used to establish typologies based on each feature. Which of the features – or which bundle of features – is chosen as a base is, at this point, still a conscious but arbitrary choice of the typology's architect. However, these suggested features of scripts will be reevaluated in the context of a functional theory of writing, where, after their relevance for processing has become clear(er), the foregrounding of one of those features for a cognitive and, in turn, informed descriptive typology of scripts might become an easier task.

Table 2: Possible categories for a script typology

| | | | |
|---|---|---|---|
| **micro-level** individual features | QUANTITATIVE | 1. number of elementary forms | How many elementary forms/segments does a basic shape consist of? |
| | | 2. density/grey-scale value | How much of the segmental space is filled with graphetic material and how much of the segmental space is empty? |
| | QUALITATIVE | 3. subsegmental structure/hierarchy | Can an internal structure be made out in a basic shape? Is there a subdivision of the segmental space in smaller spaces? (This may only be possible to evaluate in relation to other basic shapes, so it might or at least might *also* be a relational feature.) |
| | | 4. topology | How are the elementary forms/segments in a basic shape arranged? What are the relations between them? If they connect/intersect, how so? And where are they spatially positioned within the segmental space? |
| | | 5. orientation/directionality | Concerning the subsegmental structure (feature 3 above), can an "orientation" within the segmental space be recognized? For example, \|b\| would be right-oriented while \|d\| is left-oriented. |
| | | 6. intrinsic symmetry | Is a basic shape intrinsically symmetrical (whether horizontally, vertically, or point-symmetrically) such as \|A\| or \|M\|? |
| | | 7. roundness/angularity | Is a basic shape rather curved or angular in character? This depends on the elementary forms (and their hierarchy) and whether they are predominantly angular or round. |
| | | (...) | |
| **macro-level** relational features | QUANTITATIVE | 1. number of basic shapes | How many basic shapes are there in a script? |
| | | 2. number of (distinctive) features | Is there a set of visual features that act contrastively throughout all the basic shapes? And even if not, what are the distinctive features when subsets of a script (e.g. two basic shapes such as \|F\| and \|E\|) are being analyzed? In general: what are the constitutive features of the basic shapes in a script? |
| | | 3. (in)completeness | Given the features, are there any possible, well-formed basic shapes that are not yet part of the script? How many of these systematic gaps are there in a script? |
| | QUALITATIVE | 4. distinctiveness | How distinctive are the basic shapes? Is there a great deal of similarity or dissimilarity between them? |
| | | 5. visual cohesiveness, systematicity | Conversely, is the script a sound visual system? Do the basic shapes of a script have enough in common visually to be perceived as being members of the same inventory? |
| | | 6. subsegmental structure | Also (or maybe exclusively) at the relational level, the question must be asked whether there is a systematic subsegmental structure or hierarchy in a script (e.g. the hasta-coda-principle in the Roman script) that can also inform the well-formedness of (existing or non-existing) basic shapes, i.e. the (graphetic) graphotactics of a script. |
| | | (...) | |

# 2 Graphematics

The core of the present approach to graphematics is the treatment of writing systems as semiotic systems. According to the multimodular model (cf. the introduction to Part II), the graphematic module semiotically links the visual resources of the graphetic module with the linguistic resources of the language system that serves as the basis of a given writing system. It is this very semiotic relation between the visual and the linguistic that renders visual marks units of *writing* and thereby categorically distinguishes writing from other visual forms of expression such as notation or drawing. Notably, given that writing, like speech, is a modality of language, units of writing are not only linked with linguistic units but are themselves linguistic units.[68]

Notably, graphematics is mainly the subject (and actually also the result) of theoretical and descriptive reconstructions rather than a palpable phenomenon that users of writing systems themselves are aware of. Thus, graphematics is a matter relevant primarily to (grapho)linguists, whereas users are (mostly unconsciously) oriented rather towards the orthographic standard – if there is one (cf. Chapter 3). There are two main approaches to graphematics; in both, the graphematic module is claimed to contain the 'systematics' of the writing system, making it their core. In the first approach, it is assumed that the graphematic module is the sum of all (well-formed) written utterances that are possible in a given writing system – whether they are actually in use or not (cf. Neef 2005 and his *graphematic solution space*; cf. also the introduction to Part II).[69] The second approach, by contrast, depends on precisely this use and interprets graphematics mainly as the sum

**68** However, units of writing enjoy a special status: they qualify as linguistic units only by virtue of a graphematic relation, a correspondence with units of other linguistic subsystems such as phonemes and morphemes. While the latter are units of language that not only exist in every language system but are indeed constitutive of language, units of writing logically occur exclusively in languages that are equipped with writing systems – and there exist many languages without them. Thus, units of writing are logically dependent on pre-existing "linguistic units" such as phonemes and morphemes, although the nature of this dependency is a complicated matter (see below). In short, phonemes, morphemes, and the like are linguistic units of a first order and units of writing are linguistic units of a second order. Here, crucially, 'secondary' does not refer to their relevance but merely to their semiotic status. When I speak of "linguistic units" in this book, I generally refer to the former.

**69** Sometimes, even in this approach, utterances that are actually in use are distinguished from those that are merely theoretical constructs generated in accordance with the systematics (cf. Neef 2015).

of all empirical regularities (cf. Schmidt 2018). Arguably, both approaches must be combined in a comprehensive graphematics.

Terminologically, note that like *graphetics* (and designations of other linguistic subsystems such as *phonology*, *morphology*, etc.), the term *graphematics* is ambiguous: on the one hand, as mentioned above, it denotes the central module of writing systems. In turn, it is simultaneously a linguistic subsystem of languages that are equipped with a writing system. This meaning of the term is intended by Neef (2005) in the title of his book *Die Graphematik des Deutschen* ('The graphematics [= graphematic module] of German'). On the other hand, *graphematics* is also the (grapho)linguistic subdiscipline that investigates the eponymous module, which is the intended reading in the title of Fuhrhop & Peters' (2013) textbook *Einführung in die Phonologie und Graphematik* ('Introduction to phonology and graphematics'). These two meanings of the term, of course, are closely connected: the present chapter is devoted to the graphematic module of writing systems, by which it qualifies as graphematic research, or simply graphematics (cf. Berg & Evertz 2018: 188).

The tasks allocated to graphematics are captured (implicitly) by Bredel's (2008: 10, my translation) description of writing systems:

> Every writing system is equipped with a limited number of graphic material that is arranged to form motorically executable, visually sufficiently discriminable graphic units (in alphabetical systems these are letters, special characters, punctuation marks, etc.) as well as a regular combinatorics of these units (writing direction, alternation between graphic material and empty spaces, etc.) for the composition of linguistic units (such as syllables, words, sentences).[70]

As was described in the previous chapter, at first, the investigation of graph(et)ic material is pursued by graphetics. In this context, the basic shapes of a writing system's script (such as Roman script) are central. In the next step, the denotative functions of these basic shapes can be identified in a graphematic analysis, the result of which is a writing system's grapheme inventory. Notably, establishing this inventory is not entirely straightforward given that views on the nature of graphematic units (especially their relationship with language and units of language) and, in turn, their definition differ considerably. In this context, the most heated debate undoubtedly surrounds the grapheme, a concept and term that is all too often interpreted only intuitively and vaguely, by analogy with *phoneme*, as denoting the smallest, i.e. basic unit of writing. It is widely regarded as a "problematic" concept (cf. Birk 2013); it is not only fraught with myriad different attempts at a definition but also with general doubts that persist about its very *raison d'être*. The pertinent question of whether the grapheme can be reasonably conceptualized,

70 „Jedes Schriftsystem verfügt über eine begrenzte Anzahl graphischen Materials, das zu motorisch ausführbaren, visuell hinreichend diskriminierbaren graphischen Einheiten aufgebaut wird (in Alphabetschriften Buchstaben, Sonderzeichen, Interpunktionszeichen etc.), sowie über eine regelhafte Kombinatorik dieser Einheiten (Schriftrichtung, Alternation zwischen graphischem Material und Leerstellen etc.) zum Aufbau von sprachlichen Einheiten (etwa Silben, Wörter, Sätze)."

more specifically whether there can be a broad definition that holds across writing systems, is addressed in Section 2.2.

An important issue inherent in the establishment of grapheme inventories is *allography* (cf. Section 2.3), which captures the question of structural variation in writing. In short, as the prefix *allo-* that is known from designations such as *allophony* and *allomorphy* highlights, allographs are variants of the same unit. While this is the gist, the whole story is a bit more complex, as various types of allography can be identified at different grapholinguistic levels, meaning allography is not exclusively a graphematic matter. Consider, for instance, two people writing the same word by hand. The graphs in these concrete realizations are physically unique but will likely be similar visually given that in order to be recognizable to readers, they must be categorized as belonging to the same basic shapes. This type of allography is thus characterized by visual similarity and is, therefore, a graphetic matter. However, there also exist allographs that do not exhibit visual similarity but are linked by the same function; this makes their allographic status identifiable only in the course of a graphematic analysis. An example is the positional allography constituted by variants of a grapheme that occur in different positions of a word. In the writing system of Greek, |ς| occurs only at the ends of words, while the visually dissimilar |σ| appears everywhere else. They are positional allographs of the grapheme <σ/ς>. In a nutshell, this latter type of allography – graphematic allography – is necessary to distinguish the units that are individual graphemes in a writing system from those which are just variants of one grapheme.

Once graphemes and their allographs have been discovered, the system that underlies their combination must be reconstructed; by analogy with the combinatorics of other linguistic subsystems such as *phonotactics* (in phonology) and *morphotactics* (in morphology), this system is referred to as *graphotactics* (cf. Section 2.4). In addition to the combinatorial aspects mentioned in Bredel's quote, graphotactic constraints include, for instance, permissible grapheme sequences or permissible positions of graphemes within larger units such as words. Crucially, many graphotactic constraints found in the world's writing systems – such as the English rule that content words like <bee> need to consist of at least three letters (cf. Albrow 1972) – are independent of the combinatorics constraining other linguistic subsystems, most prominently phonology and morphology. That way, graphotactics highlights the necessity of studying writing systems as systems in their own right (cf. also the next section).

Finally, by mentioning composition, Bredel underlines that like other linguistic subsystems, the graphematic module is structurally and hierarchically complex. Just as, for instance, phonemes are combined to form syllables, feet, etc., and morphemes are combined to form words, graphemes are combined to form various larger graphematic units. Interestingly, for each of the linguistic units Bredel lists, analogous graphematic units have been proposed in German grapholinguistics: the graphematic syllable (cf. Fuhrhop & Buchmann 2009), the graphematic word (cf. Fuhrhop 2008), and the graphematic sentence (cf. Schmidt 2016). These need to be taken with a grain of salt, however, as they are conceived of as autonomous graphematic units, i.e. units that are defined solely on the basis of the systematics inherent in the writing system and without recourse to other linguistic subsystems. Thus, their names may be misleading – they are not merely graphematic 'depictions' of phonological syllables, morphosyntactic words, and sentences, respectively.

Crucially, these graphematic units as well as the concepts presented above – grapheme, allography, graphotactics – do not so much reveal a dependence of writing on linguistic units but instead strongly underline hierarchical and structural parallelisms between language systems and writing systems. This highlights – and quite ostentatiously so – that writing systems can be studied with many of the same methods and concepts that are used in other linguistic subfields and that, contrary to what authorities of the study of writing have claimed (cf., for example, Daniels 1991), a structural(ist) graphematics is indeed not only a possible but even a fruitful endeavor.

As already mentioned briefly in the introduction to Part II, graphematic analyses fall short when it comes to their cross-(grapho)linguistic applicability. For instance, the above-mentioned units – graphematic syllable, word, and sentence – originate from a relatively modern discussion in autonomous German graphematics. Most works treating these units state explicitly that their scope is restricted to the German – and in some cases additionally English – writing system(s) and, in turn, also to Roman script. Consequently, the conception of these units is severely limited in its application to non-alphabetic writing systems and partially even alphabets using scripts other than Roman. The fact that scripts are mentioned here already implies that some of these graphematic units actually hinge on graphetic features and are thus script-specific; this makes their definition dependent not only on graphematic but also on graphetic criteria, which are often confounded in this context, and quite detrimentally so. This, in turn, raises the question of how many of the units that are relevant across writing systems are actually determined on graphetic grounds and are, thus, script-specific. Is the diversity exhibited by the world's scripts and writing systems insurmountable from a comparative perspective? Do there even exist universal graphematic units independent of scripts or even different types of writing systems (cf. Section 2.7)? If so, how can they be defined? These questions are among those addressed in the remainder of this chapter.

To return to the larger graphematic units introduced above: their designations (such as 'graphematic syllable') are, admittedly, not entirely arbitrary. Thus, for instance, German graphematic syllables do indeed largely correspond with phonological syllables. However, the relationship between these two types of units is modeled as a correspondence and not as a dependence, and crucially, this correspondence is not even constitutive of the definition of the graphematic syllable. While this does not deny that historically or functionally, the graphematic syllable originally derives from the phonological syllable, it means that this derivative nature does not affect (or is not of interest for) methodologically autonomous graphematic analyses.

Ultimately, two separate steps can be identified as central to graphematic analyses: firstly, identifying autonomous graphematic units, i.e. units that can be discovered via an analysis of the writing system alone. Secondly, investigating whether and how these autonomous units correspond with units of other linguistic subsystems: e.g., how the graphematic syllable and the phonological syllable relate to each other. These two steps go hand in hand and are, from a global grapholinguistic perspective, incomplete without one another: identifying autonomous graphematic units without subsequently studying their place in the entire language system seems short-sighted. Conversely, establishing merely which units or sequences of writing correspond with given linguistic units but otherwise lack a *raison d'être* as functional parts of writing systems is a severe case of putting the cart before the horse. In fact, the latter may be done in fields such as phonology in which the only question of interest is how given phonemes are 'represented' by graphemes, but from a grapholinguistic perspective, any reasonable analysis must consist of either the first or – optimally – both of the steps outlined above.

## 2.1 Prologue: The relationship between language, speech, and writing

> […] contrary to received opinion ([…] in all cases misguided), the written language is not just parasitic on the spoken, it has a life of its own. This does not in any way imply, as Jacques Derrida and his more credulous epigones would have us believe, that the written language is somehow primary and the spoken, secondary; what it does imply is that the written language, in addition to or apart from slavishly representing the spoken language – obviously its primary function – has to a surprising extent its own and idiosyncratic rules, which have developed and are to some extent obeyed independently. They are, therefore, of course, all the more worthy of linguistic inquiry. (Watt 1994a: 96)

*What is the nature of the relationship between speech and writing?* This question is unavoidable and must precede any structural and functional investigation of writing, and especially of graphematics. Addressing this question explicitly is imperative since *any* analysis of writing is automatically informed by an answer to it, even if that answer often remains rather preliminary and vague (which is expected given the complexity of the matter). Simultaneously, this question is controversial and even emotionally charged in that views on it can be – and have been – instrumentalized as justifications for discounting writing as a subject unworthy of linguistic study. In any case, however, this question is most often only implicitly answered in (grapho)linguistic works (cf. Dürscheid 2016: 41). Consequently, how (grapho)linguists model the relationship between speech and writing must be reconstructed by readers on the basis of the methodology used. By comparison, when the relation between speech and writing *is* treated explicitly, unsurprisingly, as is so often the case in grapholinguistics, confusion wreaks havoc. The reason for this lies in different coexisting interpretations of various concepts relevant to this question, among them *dependency*, *autonomy*, *representation*, *reference*, and finally, *language*, *speech*, and even *writing* itself. Differing readings of these terms lead to varying conceptualizations of the relationship between writing and speech. In the following, I will attempt to present this question and to clarify some of the confusion surrounding it.

The most pressing problem with respect to this question is arguably the lack of differentiation between the two terms and concepts of *language* and *speech*. In it lies the crux of many misconceptions. For example, the equation of speech with language is the main reason for the often fervent rejection of the claim that writing represents language (cf. the 'representational myth', Stetter 2013: 89–91). While *speech* is indeed the unmarked reading of *language* (cf. Waugh 1982: 308f.; Pettersson 1996), this does not automatically imply that every time someone speaks of *language*, what they mean is *speech*. Speech and language are different phenomena and must therefore be kept apart both conceptually and terminologically. In short, language is an abstract system and speech is one of the various modalities through which language is materialized. Critical contentions – whose accuracy shall remain implicit at this point – such as "there is no language that cannot be materialized by speech (or: there is no language without speech)" (there actually is: sign language) or "there is no language if not through materialization" must be discarded, at least initially, as in the first step, it is necessary to interpret language as an entirely abstract system consisting of various subsystems: the morphological, syntactic, semantic, pragmatic, etc. The omission of phonology in this list is intentional, as the question of where and how exactly it fits into this picture is a crucial part of the answer to how speech and writing relate to one another. So, is phonology just another subsystem of language? And does phonology equal speech? These questions are seldom raised in the context of the relationship between speech and writing – but they are central.

Works treating the relationship between speech and writing often describe two opposing views that are sometimes referred to as 'hypotheses'. This terminological choice emphasizes that there cannot be a definite answer to this question, and adhering to one of the possible answers always represents a (methodological) choice (cf. Eisenberg 2006). In German grapholinguistics, the two opposing views mentioned are the *dependency hypothesis* and the *autonomy hypothesis* (cf. Glück 1987: Chapter 3; Dürscheid 2016: 35–42), which in the following I will refer to as different *views*. The autonomy view, which in essence holds that writing is autonomous from speech, is seldom interpreted absolutely, as no scholar of writing "seriously claims complete autonomy" since "written and spoken language exhibit regular correspondences on many levels of linguistic description, segmental, syllabic, and morphological" (Berg 2012: 26). A compromise is found in the intermediate, less radical *interdependency view*, which claims speech and writing exist as two relatively autonomous systems that display the aforementioned correspondences (cf. Dürscheid 2016: 35; Glück 2016a: 301f.).[71] Notably, what the two

71 In the German original, they are labeled *Dependenzhypothese*, *Autonomiehypothese*, and *Interdependenzhypothese*. However, in the literature, these hypotheses/views have

opposing views stand for – more precisely, what *dependency* and *autonomy* refer to – is also not entirely agreed on. For Glück (1987), dependency and autonomy refer to the fact that writing is dependent on or autonomous from language, while for Dürscheid (2016), they roughly mean that writing is either dependent on or autonomous from speech. When Berg (2016b: 1) posits that "the derivative nature of writing should be a hypothesis, not an axiom", it also remains implicit what exactly writing could be derivative of.[72] Concerning this very question, I argue that writing is necessarily derivative of and dependent on language, and this is not a hypothesis but, in fact, an axiom that is part of the narrow definition of writing (cf. the introduction to Part II). Before I elaborate on that, consider Figure 14, in which the three different core positions are juxtaposed.

Let us start with the view that writing is autonomous from language. According to this view, writing is interpreted as an entirely different system, in other words: as its own language. Thus, following this view, there coexist a language system that is spoken – which does not warrant equating it with speech, however, since it is still an abstract system (with morphological, syntactic, etc., subsystems) that is *materialized* by speech – and, independently of it, a distinct abstract language system that is materialized by writing (cf. Mulder 1994). In other words, speech and writing are not conceptualized as two materializations of the same language system but as materializations of two distinct language systems. Whether writing, as its own language, features the same subsystems as the other language (the "spoken one"), namely morphology, syntax, etc., remains an open question – usually, however, writing is interpreted as referring directly to extralinguistic referents (see below). Consequently, a spoken word and its written 'equivalent' are not two different materializations of the same word – instead, they are translations from one language to another. Interestingly, this does capture the actual relationship between spoken language and sign language, which is also frequently a matter of confusion.[73] Specifically, it is often erroneously assumed that sign language is a modality of the (primarily) spoken language. This is not accurate. Sign languages are independent language systems. The consensus is that they have their own 'phonology' (cf. Stokoe 1960; Sandler 2012), their own morphology, syntax, etc. An English word and its counterpart in American Sign

---

been discussed under various names, including *representational conception* (*Repräsentanzkonzeption*) for the dependency hypothesis and *distinctiveness conception* (*Distinktivitätskonzeption*) for the autonomy hypothesis (cf. Günther 1988: 72). Whether these pairs of terms are truly synonymous must be critically reevaluated, however, as *dependency* and *representation* are certainly related but not identical concepts.

**72** The article's title, *Graphemic analysis and the spoken language bias*, implies that Berg is referring to the relationship between writing and speech.

**73** Cf. Domahs & Primus (2016), a handbook dedicated to spoken, signed, and written language. Its organization implies that correlations or correspondences between speech and writing are to be treated analogously to correlations or correspondences between speech and sign language (or writing and sign language, even if the handbook completely neglects this relationship). This obscures the fact that these relations differ fundamentally (cf. Meletis 2017 and, for the same line of argument, Neef 2019).

Language (ASL), for instance, are certainly not two material shells for the same underlying linguistic representation. Instead, they are different words from different languages. Thus, while for the relationship between sign language vs. speech (as one of the realizations of a given language system such as English), the version of the autonomy view presented here is accurate, it is untenable with respect to the relationship between writing and speech. Neither from a diachronic nor from a synchronic point of view is writing its own language system.

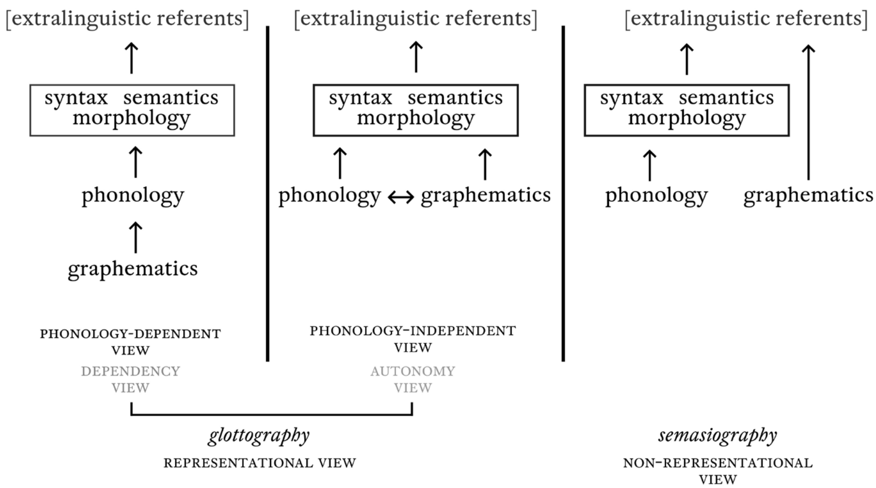



Figure 14. Different views of the relationship between writing and speech

As is evident from Figure 14, I follow Dürscheid in restricting the label *autonomy view* to refer to writing being autonomous from speech, not language. The above-rejected assumption that writing is independent of language is thus referred to as *non-representational view* (cf. Pettersson 1996). It is closely linked to the concepts of *semasiography* and *ideography*. Both are related to the idea that signs of writing refer directly to referents in the real world, i.e. extralinguistic referents, without a detour through language. A semasiographic definition of writing is given by Boone (1994: 15): “We [...] can define writing broadly as the communication of relatively specific ideas in a conventional manner by means of permanent, visible marks”.[74] Definitions like these have been referred to as *broad definitions of writing* (cf. Boone 2004: 313). Like most other scholars of writing, I reject this view and instead follow the *narrow definition of writing* which defines as writing only so-called *glottography* (literally “language writing”), i.e. writing that is related to language. Semasiography, examples of which are street signs or cave paintings, cannot be *read* – when reading is defined as the faithful decoding of a linguistic message that was encoded by a writer. Rather, semasiography – even when it rep-

**74** In Boone (2004), she critically reevaluates this definition of writing in light of the narrow definition.

resents “relatively specific ideas” – can only be *interpreted*. Ideas might be specific (however specificity is defined in this context) but the way they are (or possibly can be) verbalized is not. For example, asking a number of people what a specific street sign *means* will – provided they are familiar with its conventionalized function – yield the same or similar interpretations, but when asking them to *read* a (textless) sign, this will likely not result in the production of the same linguistic representations precisely because no fixed linguistic referents are associated with street signs. Such linguistic referents, however, are required according to the narrow definition of writing. As Daniels (2018: 157, emphasis in original) puts it, “the crucial point of this definition is that writing represents *language*”.

With that said, we return to the vagueness of the relevant terminology referred to above, a large part of which can be attributed to an inadequate or imprecise definition of writing. If writing is defined exclusively as the graphic (visual and/or tactile) representation of language, then not only is this relation axiomatic for every analysis that is based on this definition but the non-representational view must also be rejected. There is both functional and structural evidence to support this.

For functional evidence, it is helpful to return to sign language, and specifically the process of literacy acquisition in deaf people. While there exist attempts at creating notational systems for sign languages, among them HamNoSys or SignWriting (cf. Hopkins 2008; Hoffmann-Dilloway 2013, 2018), these are not classified as ‘writing systems’ as they are not used in everyday life. Consequently, what I address here under the heading of ‘literacy acquisition’ is rather the acquisition of literacy in the ‘dominant’ surrounding language, i.e. the language that is primarily spoken. Here, I only want to capture the essence of the relevant arguments; for a fuller picture, cf. Goldin-Meadow & Mayberry (2001) and Petitto et al. (2016). Many, though certainly not all deaf people in literate communities learn to read and write. This process, however, is self-evidently not as straightforward as for hearing people, which is already an indication that writing is not completely independent of speech (see below). Since the writing system that deaf people acquire is not a writing system that represents *their* native language system (which would be a system similar to the above-mentioned SignWriting and HamNoSys), learning to read and write does not mean acquiring a second modality for a language they already master but effectively amounts to second language acquisition, making the literate deaf both biliterate and bimodal. Even more so, it is the acquisition of the *secondary* modality of a second language. A native signer whose first language is ASL and whose first modality is signing, when learning to read and write in English, is acquiring a second modality (the written one) in a second language (English). Unsurprisingly, this is particularly challenging insofar as the primary modality of the language in which they learn to read and write – the spoken modality – is not, or only to a limited degree, accessible to deaf people. Mastering reading and writing, thus, is more difficult for them than for hearing people. If writing were its own system entirely and not dependent on either speech or the underlying abstract language system primarily materialized by speech, the relation between sign language and writing would equal that of speech and writing. It does not.

Another observation from literacy acquisition in deaf people strongly suggests that writing is dependent on language: deaf people who are fluent in sign language (which also cannot be taken for granted, considering, for instance, children who grow up with hearing parents) are better readers than deaf people who have neither (completely) acquired a spoken language nor a sign language. It appears that the written modality can only be successfully acquired if there is a pre-existing competence in any language system. It is the units of this language system (whether it is spoken or sign language) that the acquired written units can then be associated with. This is astonishing insofar as the written units that are learned are not directly related to the units of deaf people's first languages – and still, they link them with each other, underlining that pre-existing units of *any* language make it possible to associate the written units with *something* linguistic. This is a crucial piece of evidence that strongly indicates writing is dependent on language. However, this does not mean that writing is necessarily dependent on the specific language of which it is a modality but instead on *any* language, which suggests that writing is not a language system in and of itself.

That speech precedes writing both phylogenetically and ontogenetically is often listed as evidence supporting the claim that writing is secondary to speech. Additionally, it is noteworthy that writing does not only follow speech – it can also not be acquired without speech (or sign language, see above). Thus, writing can neither be the first language – since it is no language itself – nor the first modality of a language that someone acquires. Thus, when a child acquires English or ASL as L1, the first modality acquired will always be spoken or sign language, respectively, but not written language. Note, however, that writing can be the first (and sometimes only) modality acquired in L2 acquisition – consider Latin or Ancient Greek as examples, which are exclusively acquired through writing (cf. Dürscheid 2016: 37). The above-mentioned acquisition of the English writing system by an L1 signer of ASL is also an instance of this, as here, only the written modality of English is acquired, not (or only partially) the spoken one.

That writing is completely independent of language is questionable also from a structural point of view. First of all, if writing communicated "relatively specific ideas" instead of linguistic units, as Boone claims, this would require a sufficient number of written units to communicate all of these 'specific ideas', whose own quantity is – for lack of an operationalized definition of 'specific ideas' – difficult or even impossible to assess. With respect to large numbers of units, in fact, we know from writing system typology that there do not even exist writing systems whose basic units refer to open-class linguistic units, e.g. words, sentences, or texts since such open-ended writing systems with infinite units seem unmanageable. Also, regarding the structural makeup of writing systems, it is absurd to explain away the conspicuous isomorphy between units of writing and units of language. As will be argued below, it is very likely not coincidental that there are roughly as many graphemes in alphabets as there are phonemes in the respective languages or that in Chinese, there exists a grapheme for basically every morpheme. This isomorphy is something the non-representational hypothesis cannot explain.

After having established that writing represents language and is, thus, dependent on language, the even more important question of how writing relates to speech still remains open. To address it, I want to further specify the terminology and refer to the remaining two views – which are subsumed under the heading of *representational views* – as the *phonology-dependent* and the *phonology-independent views*, respectively (cf. Figure 14). In the following, it will be argued that three misconceptions plague the perception of the phonology-dependent view, which, for this reason, is often prematurely and inadequately rejected. These misconceptions are that a) phonology equals speech, b) all features of writing are constituted by its dependence on speech (or phonology), and c) the fact that writing is dependent on speech renders writing an object unworthy of linguistic study. Before continuing with a characterization of the phonology-dependent and the phonology-independent views, these three points will each be addressed in more detail.

a) Phonology equals speech. Speech and phonology are not the same, just like writing as a global phenomenon, a cultural technique, and graphematics, as just one (structural) part of it, are not the same. Phonology is an abstract subsystem of language that many writing systems are dependent on (to vastly varying degrees), whereas speech is the umbrella term for the acoustic materialization of language. Speech definitely includes phonology as an important part, but it is mostly phonetics and prosody[75] that are responsible for the fact that speech exhibits a number of crucial properties in which it differs from writing. For example, since in speech, the transmission of data is acoustic rather than visual, the two modalities, from a material perspective, function in completely different ways. Phonetics and, to a large degree, prosody, are commonly not represented in writing systems, as it is mainly phonological distinctions that are represented. For example, allophonic variation is usually not depicted in writing and neither are prosodic features such as accent (tone being an exception). Thus, to be answered reasonably, the question of whether writing is dependent on speech has to be restated as *Is writing dependent on phonology?* For the ensuing clarification of the relationship between writing and phonology, it helps to (roughly) conceptualize the phoneme as the smallest lexically distinctive contrast in language – bearing in mind that the definition of the phoneme, too, is a controversial matter. This working definition, of course, opens the floodgates for the next question, the issue of causation: if graphemes constitute lexically distinctive contrasts in writing, is it *because* they correspond with phonemes or do they have this function on their own?[76] This is a question that has engaged the majority of German grapholinguists in the past (cf.

75 This is accurate only if prosody is not seen as a part of phonology. Views on this differ.

76 A third possibility is that it is actually graphemes in segmental phonographic writing systems (most prominently alphabets) that allowed a segmentation of speech, which is a sound continuum (cf. Davidson 2019). In this view, the dependency is reversed and (segmental) phonology is dependent on – or even constituted by – writing.

Section 2.2). I argue that it cannot be satisfactorily answered without expanding the view to non-alphabetic writing systems. In the course of such an expansion, questions such as *If graphemes refer to morphemes in Chinese, why should they not refer to phonemes in an alphabet?* become inevitable.

b) If writing is dependent on speech (or just phonology), all features of writing must be explainable through this dependence. This is simply not true. First of all, phonology is an abstract subsystem of language and, as argued above, only a part of speech. Many of the respective features of writing and speech stem from the fact that they are distinct materializations of language. Dürscheid (2016: 24–35) lists some of the most relevant differences between them, the majority of which can be traced back to the fact that speech extends in time while writing extends in space, which corresponds with the fact that speech is continuous and writing is segmental. This makes it only logical that writing and speech each exhibit features and resources that cannot be found in the respective other. This, however, must not be used to devalue one of them (usually writing). It is not possible to directly record the volume of a voice or the emotion conveyed by it in writing, but it is equally impossible to express in speech that something is printed in italics or underlined, although, as these parallels imply, functional similarities might be found in the distinct resources of speech and writing.[77] In sum, writing is not deprived of anything when compared to speech – and vice versa.

In any case, a comparison of features that are constituted by the respective materialization is only reasonable to some degree, so the relation referred to in the title of this section can be restricted even further. In short: Writing is not dependent on speech. They are two different materializations, and only the abstract subsystem that is a part of speech is of interest: phonology. 'Writing', too, must be narrowed down further: it is not writing in general (as the global phenomenon) that is dependent on phonology, but a subsystem of it, graphematics. The answer to the question *Is writing dependent on speech?* is, thus, 'no'. However, what is often actually meant by this question can be affirmed: graphematics is dependent on phonology. However, this, too, is a generalization, as not all writing systems' graphematic modules depend on phonology (cf., for example, the morphographic kanji in Japanese). It is paramount to underline that acknowledging this dependence in no way insinuates that writing lacks idiosyncratic features that are independent of phonology and have to be analyzed independently. As proponents of the autonomy view argue correctly, graphematics' dependence on phonology only goes so far.

c) The phonology-dependent hypothesis entails that writing is a subject unworthy of linguistic study. Since some truth resides in this misconception, it proves persistent. In her characterization of the phonology-independent view (in her terminology the 'autonomy hypothesis'), Dürscheid (2016: 38,

**77** An example of this is the common interpretation of ALL CAPS as a written equivalent to shouting, which shows that there appear to be ways in which some functions of speech and writing are being compared intermodally.

my translation) notes "autonomy theoreticians argue the case for a treatment of writing as its own object of research that is to be distinguished theoretically and methodologically from speech".[78] This passage implies that grapholinguists who assume graphematics is dependent on phonology – and who stand in opposition to the "autonomy theoreticians" – believe that writing is *not* its own object of research. This view is based mainly on a number of highly prominent 20th-century linguists who are often cited as accepting only speech as a valuable and 'true' object of linguistics and simultaneously discarding writing as a mere visualization of speech without a 'life of its own'. Among them are Ferdinand de Saussure, Leonard Bloomfield, and Hermann Paul. In general, the structuralist and Neogrammarian schools are blanketly dismissed as being hostile to writing. Glück (1987: Chapter 3) offers a more fine-grained assessment of this situation and concludes that these linguists' view that "writing" is dependent on "speech" is not inseparably linked to the opinion that writing should not be studied on its own (cf. also Spitzmüller 2013: 82f.). Accordingly, I argue that it is possible and even reasonable to acknowledge a certain relation of dependence between graphematics and phonology while still treating writing as a linguistic subject and even proclaiming and supporting the establishment of grapholinguistics, an entire linguistic subdiscipline devoted to it. If the relevance that 'autonomists' vs. 'dependentialists' attach to writing as a linguistic subject is discarded as a constitutive difference between the two views, the true crucial difference between them comes to the forefront: methodology. To characterize the different underlying methodologies, let us take a closer look at the two views.

The autonomy view (cf. the middle column in Figure 14), which I specified further as the *phonology-independent view*, does not proclaim that graphematics is completely independent of phonology. What it postulates instead is that it should be studied autonomously (for an overview of the autonomy view, cf. Enderle 2005). This is captured perfectly by Eisenberg (1988: 29, my translation):

> The structural analysis of writing divorced from speech can make sense even when writing is functionally and genetically subordinate to speech in every respect. The postulate of a graphematics that is independent of phonology exists for the simple reason that because of it, a projection of the structure of speech to writing is avoided. This is necessary since otherwise structural features of writing could remain unseen.[79]

78 „Die Autonomietheoretiker plädieren dafür, die Schrift als eigenen Forschungsgegenstand anzusehen, der theoretisch und methodisch von der gesprochenen Sprache zu unterscheiden ist."

79 „Die vom Gesprochenen losgelöste strukturelle Untersuchung des Geschriebenen kann auch dann sinnvoll sein, wenn das Geschriebene dem Gesprochenen funktional und genetisch in jeder Beziehung nachgeordnet ist. Das Postulat einer von der Phonologie unabhängigen Graphematik besteht einfach deshalb, weil damit eine Projizierung der Struktur des Gesprochenen auf das Geschriebene vermieden wird. Das ist notwendig, weil sonst strukturelle Eigenschaften des Geschriebenen ungesehen bleiben könnten."

Everything Eisenberg says here is true. From this follows that the autonomy view is not actually incompatible with the dependency view – it is rather its precursor in a comprehensive and fine-grained graphematic analysis. Thus, whereas for proponents of the dependency view, written units that 'refer to' linguistic units such as phonemes constitute the smallest relevant units of writing, according to the autonomy view, the written units themselves – independent of any correspondence – are the smallest units. Accordingly, for autonomists, possible correspondences between written units and linguistic units – for instance so-called *grapheme-phoneme correspondences* – are only the product of the next, and sometimes not even necessary, step in a graphematic analysis. In short, the crucial difference between the two views lies in their respective conception of the *grapheme*, a concept that is discussed in detail in Section 2.2.

To return to the question of the distinct methodologies, in the autonomy view, a writing system's graphemes are discovered distributionally – mostly via minimal pairs. Thus, graphematics is not treated as dependent on phonology, but as parallel to it, i.e. functioning in the same way. This method of interpreting graphematics *analogously* to phonology, however, raises several new problems (discussed in Section 2.2). Ultimately, it means that in this view, too, graphematics is somewhat dependent on phonology: not structurally but methodologically. The irony in this is that proponents of the autonomy view explicitly aim to uncover features that are characteristic of writing and have no reflection in phonology or speech and strive to do so by applying to writing a methodology that was established in/for phonology. Thus, in the autonomy view, the basic units of writing – the graphemes – are defined as the smallest lexical contrasts in writing, parallel to the definition of phonemes as the smallest contrasts in speech. The description of a graphematic syllable (cf. Section 2.4) in writing systems using Roman script is another achievement of the autonomists since it is defined through mere visual – that is, writing-internal or, in their terminology, "inner-graphematic" – terms. From all of this follows that what I termed the phonology-independent view faces two challenges: 1) given its methodological dependence on the structural definition of the phoneme despite its simultaneous rejection of the relevance of phonography, it is inherently alphabetocentric. Also, 2) it heavily mixes the material (= graphetic) and linguistic (= graphematic) aspects of writing.

As this section serves only as a prologue to this chapter on graphematics, these two problems shall be touched on only briefly at this point. 1) The first problem results from the fact that the discovery procedure for graphemes is analogous to the discovery procedure for phonemes. Consequently – and ironically – it works exclusively for writing systems in which graphemes correspond with phonemes: alphabets and other segmental phonographic writing systems such as abjads and abugidas (cf. Section 2.7). For other types of writing systems – even phonographic ones such as syllabographic writing systems, let alone morphographic systems – this discovery procedure does not apply. Thus, what is termed "autonomous methodology" really is trapped in phonology's corset, after all. This results in a situation in which these conceptions regard only the smallest units of alphabets (and, presumably, other segmental phonographic systems) as so-called graphemes, whereas the smallest units of other writing systems are labeled, for example, “syllabograms” and “logograms” (cf. Glück 1987: 57). This corresponds with the autonomy view’s “axiom” that these writing systems – syllabographic and “logographic” writing systems – must “self-evidently” be analyzed in a different way than alphabets. Ironically, however, terms such as “syllabogram” and “logogram” insinuate a strong connection between the written units and the respective linguistic units they relate to (syllables and words), something that is explicitly avoided for the “grapheme”, which would accordingly and consistently have to be called “phonogram”. The question of how syllabograms and logograms can be ‘discovered’ is disregarded in autonomists’ works, as they only ever give alphabets the structural analytical treatment. The autonomy view might be best suited for isolated descriptions of individual writing systems belonging to a segmental phonographic type but reaches its limits when it comes to cross-grapholinguistic comparison or the description and explanation of universal tendencies across typologically diverse writing systems. Newer conceptions (such as Berg 2019) rely on the distribution of written units within the writing system; however, the core of this definition of the grapheme is still the claim that it is lexically distinctive. This raises the question of how such a distributional grapheme could look like in non-alphabetic writing systems. If the criterion of lexical distinctiveness cannot be upheld across writing systems, we are confronted with the questions of how different “graphemes” can be discovered in typologically diverse systems and on what grounds they can be compared, questions that will also be addressed in Section 2.2.

2) One of the autonomy view’s strongest arguments, as underlined in Eisenberg’s quote above, is that an analysis of writing in which it is treated as divorced from speech is the only reasonable (and sometimes, in fact, the only possible) method of uncovering the features that are specific to writing and have no equivalences in speech. In this context, Coulmas (1996a: 177) claims that “written language has properties not found in spoken language, and vice versa, and [...] therefore a structural description of both must precede an analysis of how sound system and writing system relate to each other”. As I argued above, however, this is true only partially. The same way it was shown that that graphematics is largely dependent on phonology (in phonographic systems), I want to suggest that graphetics, the subbranch studying the materiality of writing (cf. Chapter 1), is independent of phonetics aside from being conceptualized in analogy with it. As such, graphetics proves to be the core discipline that bundles questions concerning the modality-specific features of writing. Surprisingly, with the exception of few programmatic suggestions (cf. Günther 1990b, 1993b), graphetics is heavily underdeveloped and underrepresented in works adhering to the autonomy view. Indeed, sometimes, aspects that would have to correctly be categorized as graphetic are instead classified as graphematic. This is nowhere as obvious as in the discussion of the so-called “graphematic syllable” for German and English (cf. Section 2.4), the definition of which hinges on visual, i.e. graphetic criteria. This results in the inclusion of script-specific properties in the definition of the “graphematic syllable”, in

this case, features of the Roman script, which consequently renders the definition anything but universally applicable. This, in and of itself, is not a problem, as the description of system-specific features can of course be a valuable endeavor (cf. also Haspelmath 2010). It becomes a problem only if one takes issue with the fact that concepts such as the "graphematic syllable" are not specific to writing systems (such as German and English) but actually specific to scripts (such as, in this case, Roman). This, together with the fact that the autonomous definition of the grapheme parallels the definition of the phoneme, underlines that the autonomy view makes almost impossible a comparison of writing systems that also includes non-alphabetic writing systems.

The phonology-dependent view (cf. Figure 14: left column) also has its share of problems. The strongest characterization – and strongest weakness – of the phonology-dependent hypothesis is also found in its definition of the grapheme. What is commonly defined as a grapheme by dependentialists is the "depiction" of a phoneme. While graphemes largely correspond with phonemes in segmental phonographic writing systems, the problem of this view lies in the direction of analysis and the weighting of phonology vs. graphematics. Usually, a list of phonemes in a language's phoneme inventory is compiled and the "graphemes" depicting those phonemes are posited. It is paramount to note that this distorts the picture as it insinuates that graphemes are nothing but "visual phonemes". However, for graphematics, it is not relevant how every phoneme in a language is written. Instead, the inverse direction of analysis should be in focus: what are the phonemes that graphemes correspond with? While this directional reversal already results in a more economic endeavor, it would still lead to the assumption of graphemes such as German <ng>, since in <singen> 'to sing', this combination corresponds with, or "depicts", the phoneme /ŋ/. What is missing, thus, is a writing-specific criterion of minimality that can overwrite the criterion of "depicting a phoneme" since <n> and <g> are already individually independent graphemes of German. Such a criterion of minimality is proposed in Section 2.2. Obviously, a problem that the phonology-dependent view shares with the phonology-independent view is the focus on the phoneme. In one case, the phoneme serves as a unit of correspondence, in the other case as a model for the discovery procedure. However, despite all its problems, the phonology-dependent view is more easily extendable to other linguistic levels and is thus better suited for an application to non-alphabetic writing systems. If a core criterion for the grapheme is that it corresponds with a linguistic unit (this includes syllables and morphemes), this already highlights what diverse writing systems have in common *graphematically*, regardless of how distinct they are graphetically (see above).

Dürscheid (2016: 36) lists two functional arguments in favor of the view that writing is logically dependent on speech: that there exists no language that is equipped with a written modality but lacks speech[80] and that writing is secondary to speech both phylogenetically and ontogenetically. These are generally

80 This should be extended to "there is no language with writing but without speech or

agreed upon by both dependentialists and autonomists (cf. Birk 2013). Thus, the difference between these views is more fine-grained and located, as we have seen, at the methodological and structural levels. In a nutshell, the autonomists argue that writing's logical dependence on speech does *not* entail structural dependence, resulting in idiosyncratic features of writing that can only be discovered with an autonomous methodology. However, a more universal look at this question that takes into account also non-alphabetic writing systems raises serious challenges for this view of structural independence. An alternative solution to this problem that was not yet mentioned is that generalization is impossible because segmental phonographic writing systems have a different status than non-segmental ones. The key lies in the two core problems introduced above; they will be addressed in more detail in the following sections: the autonomous, analogous definition of the grapheme (Section 2.2) and the definition of a graphematic syllable (Section 2.4). At the end of this chapter, an epilogue will bring together the findings of the subsections and draw conclusions that shed further light on the question of how writing and speech relate to each other.

## 2.2 Grapheme[81]

Every writing system is structurally complex. In every writing system, regardless of how (much) it differs from other writing systems, there exists some kind of minimal unit with a linguistic function that is used compositionally to build larger units. Irrespective of which exact functions this minimal unit fulfills, it represents the basic entity of writing systems. In analogy to basic linguistic units such as the phoneme or the morpheme, it has been proposed that it be called *grapheme*. So far, so good. A number of scholars across disciplines have used this term but often have done so vaguely, failing to make explicit what they mean by it.[82] The instances in which *grapheme* is indeed defined reveal that understandings of the concept diverge considerably, reflecting a lack of consensus over what the term signifies.

More than three decades ago, Manfred Kohrt (1986: 81), who had authored an extensive treatment of the grapheme (Kohrt 1985), noted: "[P]erhaps you are convinced that the problem of providing [the grapheme] with an adequate definition has already been solved or is simply insignificant" and shortly thereafter concluded: "As for me, I share neither of these views". To this day, the matter

signing", as not only speech but also signing can serve as first modalities of a language and as a basis for writing.

81 This section is largely based on Meletis (2019).

82 In this context, Mugdan (1990: 50, emphasis in original) comments on the problematic vagueness of blanketly coining linguistic terms ending in *-eme*: "The insatiable need for designations of linguistic units ensured the remarkable success of the new derivational pattern, but the combinations of *-eme* with suitable roots (*morph-*, *graph-* and the like) have such general meanings ('unit of form', 'unit of writing' etc.) that one could utilize them to name any of several different concepts".

remains relevant but unsettled. Despite this, there currently appear to be no aspirations to negotiate a definition of *grapheme* that researchers interested in matters of writing would agree on. Quite to the contrary: the terminological and conceptual chaos surrounding the grapheme has led some eminent scholars of writing to reject the concept altogether (see below), resulting in a depreciation and, consequently, a palpable stagnation of any fruitful discussion surrounding it.

Instead of *grapheme*, more or less established terms such as *letter*, *character*, or vague placeholders like *sign (of writing)*, and *symbol* are currently in use. They all lead double lives as lay terms and quasi-technical terms, and their use is problematic. For example, labeling the smallest functional unit of the German writing system 'a letter of German' might be terminologically adequate in an isolated description of the German writing system. It might even allow for comparisons with other writing systems of the same type, i.e. alphabets, and comparisons of their units, i.e. the alphabets' respective 'letters'. Ultimately, however, this conceptual and terminological choice dissociates the results of such an alphabet-specific description from a global theory of writing by obscuring the fact that 'a letter of German' may share features with 'a character of Chinese' that allow for a unified classification of these units as graphemes of their respective writing systems. In general, comparative analyses across various types of writing systems become more difficult without a shared conceptual and terminological framework, which might be the main reason why they are so rare.[83] Comparison, however, as is argued in this book, is undeniably the crux of an emerging theory of writing, and in order to compare diverse writing systems, we are in need of a benchmark. Therefore, positing a concept of *grapheme* and defining it in a broad manner that makes it applicable to all writing systems and facilitates comparisons between them is certainly an important and promising endeavor for the advancement of grapholinguistics.

In this section, first, the most important views on the grapheme will be presented, and on this basis, a new and improved definition of the concept will be proposed. Crucially, unlike past alphabetocentric definitions, the suggested conception of the grapheme takes into consideration all types of writing systems, which will be demonstrated by examples from a variety of systems.

### 2.2.1 Different conceptions of *grapheme*

The term *grapheme* made its first appearance at the beginning of the 20th century. Jan Baudouin de Courtenay is commonly credited with coining it[84] and interpret-

83 But see, for example, Häffner's (2009) treatment of the German and Japanese writing systems within one theoretical framework.

84 Common consensus is that the term *grapheme* was coined independently at least twice: in addition to Baudouin de Courtenay, Kohrt (1986: 82–83) mentions Finnish linguist Aarni Penttilä as having introduced the term – albeit in a different sense – in

ing it as the psychological representation of 'letters' (cf. Kohrt 1985: 171–172 for a list of differing but conceptually related uses of the term in Baudouin de Courtenay's works). Crucially, early definitions of the concept were almost always linked to works of phonology, and "[e]very conception of the 'grapheme' had some previous interpretation of the term 'phoneme' which served as a model" (Kohrt 1986: 82). The concept's further development moved into two separate directions closely linked to the respective grapholinguistic tradition in question: in Anglo-American research on writing systems, the term was largely abandoned, and it is absent in milestone works such as Gelb's (1952) *A study of writing*. Notably, in some prominent textbooks and introductions pertinent to grapholinguistics, the grapheme is also attended to very briefly and unspecifically (Sampson 2015) or not mentioned at all (Coulmas 2003). By contrast, in the German-speaking tradition of research on writing, the grapheme quickly advanced to a central subject of discussion when German grapholinguistics flourished in the 1980s; it was eventually joined by proposals of other graphematic units such as the graphematic syllable (Fuhrhop & Buchmann 2009), the graphematic foot (Evertz 2018), the graphematic word (Fuhrhop 2008), and the graphematic sentence (Schmidt 2016), which are treated in the following sections of this chapter.

Overall, four major strategies of dealing with the grapheme have emerged: rejecting the grapheme, not defining the grapheme, preferring the concept of *letter* over *grapheme*, and, most importantly, assuming and defining a concept of *grapheme*. While the first two strategies are characteristic of the Anglo-American literature, the latter two are closely associated with German grapholinguistics. In the following, each of them will be considered in turn.

The first strategy is (1) the rejection of the grapheme, and the most prominent scholar adhering to it is Peter T. Daniels, who is simultaneously also the best-known scholar of writing systems. In one of his more recent encyclopedia entries on writing systems, under the heading 'Writing and Language', Daniels lists various reasons that lead him to believe that the grapheme "has become nothing more than a pre-theoretic, fancy, scientific-sounding word for 'letter' or 'character' and ought not to be part of technical discourse" (Daniels 2017: 88). I want to discuss his points one by one to show that opposing and thereby dismissing the grapheme is, for the sake of a broader grapholinguistics and a vital comparison of writing systems that transcends mere individual descriptions, not the way to go.

Daniels' main point traces back to his rejection of a 'structural graphemics' (Daniels 1991) in which he argued writing and language must be treated differently because language is unconscious while writing is conscious.[85] His rea-

1932. Furthermore, Daniels (2018: 171) claims that in the 1930s, Benjamin Lee Whorf, W. Freeman Twaddell, and R. H. Stetson also all used the term *grapheme* independently of one another.

85 In this respect, Kohrt (1986: 93) agrees with Daniels, stating: "Writing always presumes some kind of consciousness (which is not necessary in the sphere of spoken language) […]". Note, however, that Kohrt speaks of spoken language while Daniels speaks of language in general.

soning is not primarily linguistic – as he claims, "writing is not like language, and it is not like language for biological reasons" (Daniels 2017: 88). Observations that he cites to underline this are that children do not acquire literacy without instruction and that there appears to be no critical period for it since illiterate adults can still learn to write and read without difficulty. Additionally, he notes – much like Dehaene (2009) famously formulated – that there is no "special capacity" (Daniels 2017: 88) for writing because it is too recent an invention for humans to have evolutionally adapted to it.

Furthermore, Daniels states that phonemes and morphemes are unconscious properties of language "and other realms of human behavior" (Daniels 2017: 88), which also accounts for the suffix *-eme* in the designations of those concepts that is used to mark abstractions. Writing, however, is "not an unconscious, built-in feature of a mind" and therefore "cannot a priori be assumed to be analyzeable in a parallel way" (Daniels 2017: 88). This seems to be Daniels' only – but nonetheless major – contention against *graphem(at)ics* as a subbranch of linguistics in which writing systems are studied as linguistic systems, including the units that have been assumed in analogy with other linguistic units such as the phoneme or the morpheme. The specific reservations he expresses against the grapheme – in a nutshell, its seeming multifunctionality *within* individual systems and the different functions it supposedly fulfills *across* systems – will be discussed below, but first, the claim that writing cannot be studied analogously to language warrants further consideration.

There really is no clear-cut distinction between 'writing and language', since writing – if defined glottographically, which nowadays is the broad consensus in writing research – is always a form of language, specifically a modality of a language system that complements the (primary) spoken modality. Logically, thus, it is not positioned at the same level as phonology or morphology, subsystems that are necessarily *inherent* to each language system. Indeed, phonology and morphology are not optional in a language system. By contrast, graphematics is. And still, it is striking how graphematics behaves much like these other systems in various respects: as established above, what is central for the proposal of a grapheme definition is the structurally complex nature of graphematics.

Most statements Daniels makes about writing are correct. It is not acquired without instruction and it is, in fact, both phylogenetically and ontogenetically secondary to speech (cf. Dürscheid 2016: 30). There is no brain region that has evolved specifically for reading or writing as it is rather other areas with similar functions that have been neuronally 'recycled' in order to take on the tasks of reading and writing (cf. Dehaene 2009: 144–147). And finally, writing is more conscious than language as writers and readers need to be – at least in the process of acquiring a writing system – consciously aware of the units of writing they employ. And it is even more than that: readers and writers are not only conscious of the units of writing, but in using them, they must also be (or become) aware of the linguistic units (or properties, as per Daniels) that the units of writing relate to. Writing not only requires phonological or morphological awareness, it more generally requires metalinguistic awareness. Writing not only is conscious itself – it also makes us conscious of language.

Ultimately, this means that writing, as an optional modality of language, always constitutes an analysis of language. As Reiner (2000: 1) puts it: "The invention of writing may well be considered testimony of linguistic analysis of the spoken language", with phonology, morphology, syntax, etc. being linguistic levels that are subject to this analysis. Interestingly, even Daniels (2013: 53) himself has stated something similar, namely that "[e]very writing system represents a 'native-speaker analysis' of a language, and as such at every stage of its development it reflects what its users consciously 'know' about their language" (cf. also Dürscheid 2016: 39).

The question, now, is whether the fact that writing operates at another level, making the 'unconscious' properties of language conscious, should preclude us from analyzing it with the same or at least similar tools and concepts that we use to analyze other levels (such as phonology and morphology). Rogers (2005: 11) negates this by claiming that "the fact that the data of language and writing are different in nature does not preclude our using a similar theoretical framework" (cf. also Primus 2004: 237). The most crucial reservation about Daniels' rejection of graphematics cannot easily be resolved: is writing really always so conscious? Once acquired successfully, are the processes of reading and writing not often dominated by automatisms? If someone is a fluent and competent writer in his native (or even a second) writing system, is there really a conscious process involved in writing down an item on a shopping list? If I want to write down 'milk', do I always map each phoneme that I intend to write to a grapheme and, in a next step, to a basic shape that I must then graphomotorically substantiate in the form of concrete graphs? Conversely, is pronouncing something in speech always unconscious? And with respect to performance, a question that challenges the claim that writing is always conscious is how and why unconscious mistakes happen in writing.

In any case, we must carefully separate the discussion of the grapheme as a unit relevant in processing on the one hand, which is addressed by this question of consciousness, and the assumption of the grapheme merely as a unit of description on the other, which remains unaffected by this question. The present section is focused on the latter, i.e. the grapheme as a descriptive unit. Whether such a unit is psychologically real must be evaluated in a separate step.

With respect to the distinction between an etic and an emic level, which goes back to Kenneth L. Pike ([1954] 1967: Chapter 2), it must be underlined that there is undoubtedly an emic level in writing: concrete substantiations can be – or even must be – classified into abstract categories to make an analysis even possible. Fittingly, an emic unit is defined as "an invariant form obtained from the reduction of a class of variant forms to a limited number of abstract units" (Nöth 1990: 183). As will be shown below, graphemes are the smallest emic units of graphematics: they relate a graphetic form, a basic shape, itself an emic unit (of, ironically, graphetics, cf. Section 1.2.1), to a linguistic unit.

Daniels mentions two more arguments against the grapheme that go hand in hand. First, he claims that what has been called *grapheme* in different works has varying functions, depending on the unit that it was analogously modeled after: like the phoneme in a phoneme inventory, it sometimes is seen as a unit of a set of units that comprise a writing system; like the tagmeme, which itself is defined as "the correlation of a grammatical function or slot with a class of mutually substitutable items occurring in that slot" (Elson & Pickett 1962: 57), it is the correlation between a syntagmatic function and a paradigmatic filler; like the morpheme, it is "a minimal extent of something" (Daniels 2017: 88). While he agrees on the fact that each of these definitions might be suitable for given (types of) writing systems, Daniels notes that they cannot be satisfactorily reconciled in a coherent definition of the grapheme that applies to all writing systems – something that, however, is demanded from a universal grapheme definition.

In a different contribution, Daniels, together with his collaborator, psychologist David Share, claims that "'[g]rapheme' has had so many different interpretations that in writing systems theory it is meaningless" (Share & Daniels 2016: 23). While the first part of this statement is undeniably true, as will become evident below, it cannot be seen as a sufficient reason to altogether dismiss the idea behind the term or even the term itself and to discard it as "meaningless". As Rogers (2005: 11) posited with respect to the grapheme (and other grapholinguistic terminology), "we can define and use our terms carefully". One of the many definitions of the grapheme that have been proposed might be the right one, or the right one has yet to be formulated. The truth, I believe, lies somewhere in the middle: some of what has been proposed under the heading of *grapheme* will turn out to be accurate while other aspects can indeed be dismissed. What is at the core of Daniels' problem with the grapheme is that the shocking majority of efforts in defining it has taken place outside the larger linguistic picture. While in the history of the phoneme and the morpheme, definitions necessarily ventured for universality pretty quickly, in the case of the grapheme, it was always 'a grapheme of German' or 'a grapheme of Chinese', but never a 'general grapheme' that can truly explain how writing in general functions at its core. This, of course, is an understandable reaction to the typological diversity of writing systems (cf. Section 2.7), which, at first glance, might discourage scholars of writing from seeking out the possibly universal nature of the grapheme.

Another aspect that Daniels (2017: 88) briefly discusses is that "many alphabets use a pairing of symbols – capitals and lowercase, majuscule and minuscule – that has no equivalent in sound systems". In the same vein, he mentions italics (cf. Daniels 2018: 169). While it is true that neither capitalization nor italics have direct equivalents in speech, this does not necessarily translate to an argument against analyzing the graphematics of a writing system in a structuralist way. It also does not negate the possibility of defining a grapheme. A more adequate reaction to these observations is to acknowledge that while writing represents language in a written form, it *also* has ‘a life of its own’ (cf. Section 2.1). In other words, not everything in writing must refer to something in “sound systems” and not everything in writing must have an “equivalent in sound systems”. Writing is a system in its own right, not least because of its visual materialization that exhibits altogether different regularities than the acoustic materialization of speech (cf. Chapter 1). To return to Daniels’ examples: it is accurate that the uppercase and lowercase ‘letters’ of an alphabet do not refer to different phonemes, meaning the difference between them reflects no difference at the phonological level, and there is also no functional equivalent to this case distinction in phonology, i.e. “lowercase and uppercase phonemes”. However, this case distinction in writing has a different function at a different level (see below), e.g. the morphological level (by differentiating parts of speech) or the syntactic level (by marking the head of a noun phrase as in German), or even the graphematic level itself (by marking the beginning of a graphematic sentence). Ultimately, a treatment of writing with linguistic methods does not require that a) everything in writing must be derived from phonology (or “sound systems”), which is what the *phonology-dependent view* holds (see above) or that b) everything in writing must have an (analogous) equivalent in phonology or some other linguistic level (*phonology-independent view*). Both are partially true but neither of them gets it completely right – and even in combination, they do not capture the complexity of the whole picture.

Of Daniels’ arguments against the grapheme, I want to return to the one lamenting ‘too many incoherent definitions’ as it serves as a transition to the second strategy of dealing with the grapheme. In many works, scholars (2) use the term *grapheme* but fail to define what they mean by it. Instead, they use the term in a vague manner,[86] ultimately leaving it up to readers to decipher its meaning. In doing that, they avoid committing to a specific definition of the term (which, as this section proves, is indeed a difficult affair). This empty and rather careless use is indeed so frequent that Daniels’ aversion to the grapheme becomes largely relatable. However, this type of use is predominantly found in non-linguistic works, e.g. in psychology, the cognitive sciences, pedagogy, etc. Alas, because of the interdisciplinarity of grapholinguistics, the misuse in these works affects the whole field. Moreover, it is not only publications on alphabetic writing systems that tend to (mis)use *grapheme* but also works on non-alphabetic writing systems.

---

86 At this point, let me recount an anecdote from a workshop on writing systems I attended in Nagoya, Japan, in 2017 (the 11th international workshop of the *Association of Written Language and Literacy).* There, I presented essentially the content of this section on the definition and the use of the term *grapheme*. I was the third speaker on the workshop’s first day, so most of the other talks followed mine. What my talk ignited was rather entertaining: some speakers who presented after me used the term *grapheme* in their talks only to instantly become aware of what they had done and back away from the term right after. The laughter of the crowd as well as the presenters themselves to comments such as “I am afraid to use the term *grapheme* now” or “I suppose I am using *grapheme* wrong” proved that very frequently, there is no reflection on what *grapheme* really stands for and that the term is used in various meanings.

Take, for example, a study about handwritten character production in Chinese (cf. Chen & Cherng 2013). In the first paragraph of the paper, the authors state "[t]he letters or graphemes serve as functional units in the orthography of a word" (Chen & Cherng 2013: 1). Here, evidently, *grapheme* is used as a synonym for *letter*, corroborating Daniels' claim that the term often serves as a more scientific-sounding alternative that (in this case) could be abandoned without loss.[87] Next, in a study about reading in Thai, Winskel & Iemwanthong (2010: 1024, 1028) mention "consistent grapheme to phoneme mapping" or "children's phonological knowledge and ability to map sublexical units onto graphemes". However, they fail to define what a grapheme is or what units comprise the grapheme inventory of Thai. *Grapheme* is obviously meant to be some sort of functional written unit, but in a typologically phonographic but non-alphabetic writing system such as Thai, is it so intuitively obvious what graphemes are? Is it just the consonant 'signs' or also the secondary and dependent vowel 'signs'? Winskel & Iemwanthong (2010) also frequently use the term as a part of the common phrases "phoneme-grapheme-correspondences" and "grapheme-phoneme-correspondences". These technical expressions have their roots in works on alphabets and are not without flaws, especially if used without a prior definition of the individual components. Taha (2013), in a study on reading and spelling in Arabic, also addresses these correspondences. When he first mentions basic shapes (only scriptual basic shapes, i.e. no digits, special characters, or punctuation marks, cf. Section 1.2.1), he uses the term *letter*: "Arabic is a language written in an alphabetic system of 29 letters [...]. Arabic letters have more than one written form, depending on the letter's place in a word: beginning, middle, or end" (Taha 2013: 725). Here, *letter* seems to be an abstract notion whose materialization are *written forms* that depend on the position of the letter. In a later passage, however, Taha changes his terminology, writing that "the basic and non-connected shape of the grapheme س could change according to its placement within the word" (Taha 2013: 725) – letters are now *graphemes*, written forms are now *shapes*. Uses such as the above are numerous. They contribute to the overall impression that indeed, *grapheme* has no fixed meaning.

To turn to the third strategy, we must consider approaches in which the grapheme is (3) not dismissed altogether but avoided (or marginalized) in favor of other (supposedly) more specific terms.[88] A scholar adopting this strategy is Martin Neef, who like Daniels bemoans the fact that the grapheme has been used with a variety of heterogeneous meanings (cf. Neef 2005: 36). In his seminal work on the graphematics of German (Neef 2005), Neef opts out of using *grapheme* and

87 Additionally, they describe units called *logographemes*, "the smallest units in a character that are spatially separated, and they appear in many characters" (Chen & Cherng 2013: 2).

88 Berg's study serves as a fitting example of this strategy. He acknowledges graphemes but does not exactly deal with them: "This paper will primarily deal with letters, not with graphemes. The distributional analysis is facilitated if we deal with letters first. This is not to say that graphemes can or should not be part of either method, but at this stage of investigation it seems legitimate to exclude graphemes" (Berg 2012: 39).

uses *letter* (German *Buchstabe*) instead. However, unlike Chen and Cherng, he does not treat the terms as synonyms, and unlike Daniels, he does not argue that there is no (or that there cannot be a) unit called *grapheme*. However, he claims that the concept of letter is sufficient in the description of the German writing system and that, while it might be a useful concept elsewhere, the grapheme is superfluous. Especially in the present context of a comparison of typologically diverse writing systems, the choice of the letter as the smallest functional unit must be critically reevaluated – even if it was proposed within a system-specific analysis of German.

Neef (2005: 37) uses the German graphematic word <Schnee> 'snow' to illustrate that the grapheme is a different concept than the letter. Since he analyzes <sch> as one grapheme (following both the referential and the analogical views, with exceptions, see below), <Schnee> consists of the four graphemes <sch>, <n>, and two instances of <e>, and simultaneously of six letters, because <sch> is itself a complex grapheme made up of the three letters <s>, <c>, and <h>. Since the constituents of <sch>, <s>, <h>, as well as possibly <ch>, are also commonly analyzed as graphemes, Neef notes that in any theoretical framework that includes graphemes, the grapheme must be defined in a way that accounts for the possibility that it is comprised of one letter *or* a combination of multiple letters. He concludes that a theory of graphematics can function either with the unit letter alone or with the units letter *and* grapheme. However, due to complex graphemes such as <sch>, it cannot function with the grapheme alone.[89] Ultimately, for reasons of economy, Neef decides that the letter suffices for his graphematic theory of German.

Notably, this leaves open the question of how complex units such as <sch> are to be treated. Since the basic unit is the letter, Neef chooses to call these units *fixed letter combinations* (German *feste Buchstabenverbindungen*, cf. Neef 2005: 41). Furthermore, because *letter* already functions as the designation of the smallest abstract functional unit, Neef is in need of a supplementary term to designate the (also abstract) graphetic manifestation of these abstract units, what I have called *basic shape*. Here, he chooses *letter body* or *letter gestalt* (translated from German *Buchstabenkörper*, cf. Neef 2005: 39).

Conceptually, I agree with Neef on all accounts. I merely propose a terminological shift. My insistence to keep *grapheme* is based on the fact that it allows (or facilitates) comparisons between different writing systems. While *letter* or *character* might suffice in individual descriptions, a definition of *grapheme* that captures the minimal functional unit of *any* given writing system reflects that at the core, they all share a crucial function: visually representing language (and not just speech). As mentioned above, Neef's introduction of *letter body* illustrates

89 „Unabhängig von der Frage einer intensionalen Definition des Begriffs Graphem liegt nach meinem Verständnis in einer Theorie nur dann die Grundeinheit Graphem vor, wenn Grapheme sowohl aus einem als auch aus mehreren Buchstaben bestehen können. Auf der Basis dieser Überlegungen kann eine Theorie der Graphematik entweder nur mit Buchstaben oder mit Buchstaben und Graphemen arbeiten, nicht aber mit Graphemen allein" (Neef 2005: 38).

that his functional interpretation of *letter* leaves a vacancy regarding the graphetic units that take part in graphematic relations. I had argued that *letter* is, in fact, a graphetic term, as is *character* (cf. Section 1.2.1). Both terms designate types of basic shapes and can be replaced by the umbrella term *basic shape*. The additional information they provide is what kind of script they stem from, although that is only partially true for *letter*, which at this point enjoys great popularity whenever talk is of the units[90] of an alphabet – not only alphabets using Roman script but also the Georgian, Armenian, Cyrillic, and Greek scripts. Furthermore, the term has spread to the basic shapes of scripts that are used for typologically different writing systems: take Arabic, whose units are also sometimes referred to as *letters* (cf. Saiegh-Haddad & Henkin-Roitfarb 2014: 15). Ironically, this overgeneralization of the term reveals aspirations of finding a common ground across different writing systems by using the same designation for their units. It is this very motivation that also justifies a unified definition of the grapheme.

If, as in the approach proposed here, *letter* is interpreted as a script-specific term for *basic shape*, i.e. a graphetic unit, there remains a need for a label designating the basic graphematic unit that can be found cross-grapholinguistically. If this unit is termed *grapheme*, combinations of graphemes are called *complex graphemes.* In a nutshell, Neef's terms *letter body, letter,* and *fixed letter combination* are replaced with *basic shape, grapheme,* and *complex grapheme*, respectively.

Another model that does not dispense with the grapheme but, in comparison with older approaches, dramatically reduces its relevance is the suprasegmental model developed by Beatrice Primus and her colleagues (cf. Figure 15). It evolved from the idea that the syllable can be described not only as a salient unit of spoken language but of sign language and written language as well. As such, it is interpreted as a unit independent of medium (or rather modality) (cf. Primus 2003; Section 2.4). Like in Neef's model, in this model, too, the *letter* is interpreted as an abstract linguistic unit that fulfills the most crucial function (cf. Berg 2019: Chapter 3.1). Whereas for Neef, letters are abstract functional units, and their concrete physical manifestations, the letter bodies, are analyzed holistically (and are interpreted as arbitrary[91]), Primus also proposes a level that is located below the level of letters. According to Primus (2004, 2006), letters of the Roman 'alphabet' derive their referential value and, consequently, their function compositionally on the basis that the visual features of their elementary forms correlate with phonological features. That letters are not interpreted merely as arbitrary holistic shapes but as units that are simultaneously visual and functional and exhibit a complex internal structure suggests that graphetics and graphematics are more tightly interlocked than is traditionally assumed.

---

**90** Here, I speak of units and not of basic shapes because much like *grapheme*, *letter* has often been used vaguely; thus, it is not always clear whether what is meant by it is the visual form, its linguistic function, or both.

**91** However, Neef (2005: 39) does not rule out the possibility that letter bodies can be structured complexly. In his approach, however, this hypothetical complexity does not play a role.

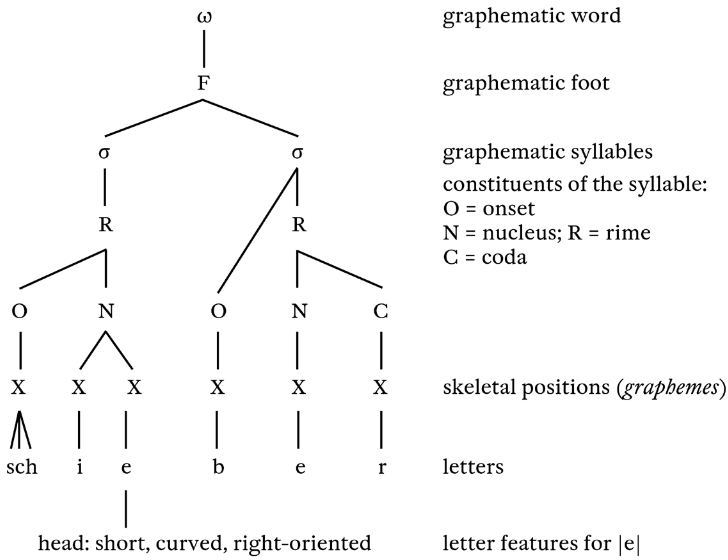



FIGURE 15. Suprasegmental model of writing, from Berg, Primus & Wagner (2016: 351)

A second aspect that distinguishes this suprasegmental model from Neef's model is that an additional level is assumed above the level of letters, the level of skeletal positions. The authors suggest that these skeletal positions could possibly be graphemes.[92] In most cases, one letter will be associated with one skeletal position, meaning single letters for the most part simultaneously function as graphemes. By giving the term and the concept *letter* priority, Primus and her colleagues imply – similarly to Neef – that because letters so often coincide with graphemes, they actually do suffice. The concept of grapheme, by contrast, is only relevant for letter combinations that behave like single letters, for example, German <sch> (cf. Schmidt 2018: 138). However, letters can also be combined and a combination of multiple letters can be associated with a single skeletal position, which is how in this model, too, <sch> is regarded as one grapheme consisting of three letters (cf. Berg, Primus & Wagner 2016: 351). Unlike Neef's model, in which these sequences are termed fixed letter combinations, the suprasegmental model chooses *grapheme* as a designation for these complex units.

92 Schmidt (2018: 128) points out a problem of treating skeletal positions as graphemes: in graphematic words such as German <beten> 'pray' or <lesen> 'read', the first vowel, in each case an instance of <e>, would be associated with two skeletal positions, respectively, because these graphematic words are simultaneously (as graphematic feet) trochees, meaning their first graphematic syllables are 'strong' or 'prominent'. As Schmidt argues, it would be absurd to claim that the first instances of <e> in these words are two graphemes, respectively. Skeletal positions, thus, do not always correspond with graphemes. Schmidt also assumes that Primus and colleagues silently acknowledged this by changing the skeletal positions' designations from 'G' (for grapheme) to the more neutral 'X' in recent versions of the model.

To sum up, the reliance on the letter as the most central unit in these otherwise promising approaches is problematic because it obscures the fact that non-alphabetic systems might share features with alphabetic systems (cf. Meletis 2017: 112).[93]

What remains is the fourth and clearly most relevant strategy of dealing with the grapheme: (4) the assumption and acceptance of graphemes as relevant functional units of writing systems. In the models and theories of the grapheme's proponents, it plays a crucial role. However, even among them, there is no consensus about how it is identified and what functions it fulfills in a writing system as well as the language system as a whole. The two most influential conceptions of the grapheme are based on the (4a) *referential view* and the (4b) *analogical view* (cf. Kohrt 1986; Lockwood 2001). These are closely linked to the differing views of the relation between writing and language – or more narrowly, writing and speech. In Section 2.1, these were termed the i) *phonology-dependent view* and the ii) *phonology-independent view*, respectively. While scholars who adhere to the first of those views interpret writing as a representation of speech (or phonology), proponents of the latter view conceive of writing as a form or *modality* of language that is (largely) autonomous from speech (or phonology). What, now, do these different views mean for the definition of the grapheme?

In the referential (or *representationalistic*, cf. Coulmas 1996a: 175) view, graphemes are signs of phonemes, while in the analogical (or *distributionalistic*, Coulmas 1996a: 175) view, graphemes are identified by analogy with how phonemes are 'discovered' and are, thus, defined as the smallest functional units of writing regardless of whether they correspond with phonemes (or other units of language). Like Kohrt (1986), I argue that individually, both views prove insufficient for explaining what the grapheme is – at least cross-linguistically. However, I also agree with Lockwood (2001: 307) in that "students of writing need to include both kinds of relations in any model they adopt". Accordingly, there is some truth in both views but neither of them gets it completely right. The obvious Eurocentrism that plagues both conceptions will be addressed after elaborating on the challenges they face even when applied exclusively to alphabets, the very systems they were designed for.

Proponents of the referential view treat graphemes as written units that stand for phonemes, i.e. units that function as "phoneme signs" (Kohrt 1986: 84). It is confronted with several problems. As Günther (1988: 76) points out, if graphemes are derived from phonemes, there is no need for the concept of grapheme to begin with, as they would simply be written labels for phonemes. In this

93 Focusing on the letter might be understandable and acceptable in the context of descriptions of alphabets but not in the context of an attempt to establish a general model of writing (systems). Thus, choosing *letter* as the unit of writing for the title of a handbook – as in *Laut, Gebärde, Buchstabe* (transl. *Sound, gesture, letter*) (cf. Domahs & Primus 2016) – is questionable (and a missed opportunity on top of that), especially since the selected units for speech and sign language are applicable universally.

sense, graphemes would not be units but relations, correspondences, or rules.[94] Another issue that Günther addresses is allography. In the referential view, allographs are conceived of as variants of writing one phoneme, meaning allographs are assigned to phonemes instead of graphemes (cf. Günther 1988: 76; cf. Section 2.3.4). For example, Garbe (1985: 12f.), who is harshly criticized by Günther, lists <f ff v w + fe ph> as allographs of the phoneme /f/ in German, with <fiel/Suff/viel/Möwchen + Safe/Philister> 'he fell/boozing/much/small seagull + safe/philistine' provided as examples of words in which these allographs occur and correspond with /f/ (cf. also Zifonun et al. 1997: 273–280 for an exhaustive list of such German sound-grapheme correspondences).

This unveils two further problematic aspects of the referential view: firstly, the direction of analysis is 'from phoneme to grapheme'. Primacy is thus given to encoding phonology in graphic form, i.e. to production processes. This is not a problem *per se*, but a choice – a choice that, however, through prioritizing phonology, yet again gives primacy to speech and strongly insinuates that writing is but a derivation of speech. Modern approaches such as Neef's (2005) either choose the opposite direction for their analysis or take both into account and postulate bidirectional mapping relations (cf. Evertz 2016: 381).[95] Secondly, Kohrt's (1986: 87f.) major criticism of the referential view echoes Günther's sentiment that graphemes are not units but relations: phonemes are subject to context-sensitive variation. Accordingly, if there were a one-to-one equivalence between phonemes and graphemes, which is what "has always been considered as something that should be aimed at" (Kohrt 1986: 87), writing systems would not be usable tools of written communication but transcriptions. They are not. Specifically, Kohrt (1986: 88) writes

> what you will get [...] is a phonemic transcription [...] – but you will never arrive at something like a traditional orthography. [...] As far as orthographic writing is concerned, the structure-determined phonemes cannot serve as the relevant correspondence units; it would be a mistake to look at the written marks that are important for the orthography as 'graphemes' which are defined by their relation to the phonemic entities.

A variant like <ff> is, thus, not (as Garbe proposes), an allograph of a phoneme /f/ (or even of a grapheme <f>). Since graphemes are to be defined as *minimal* functional units of writing, <ff> is in any case a sequence of two instances of the grapheme <f>. The doubling is determined by graphematically higher levels (cf. Birk 2013; Schmidt 2018: 32f.), in this case, the syllabic level: in <Löffel> 'spoon', for

94 „Der Begriff Graphem in der Lesart der Repräsentanzkonzeption ist kein Analogon zum Begriff Phonem, das solchermaßen definierte Graphem ist systematisch nicht als Einheit klassifizierbar. Es ist eine Relation, die Beschreibung eines Umsetzungsprozesses, eine Regel oder was immer, jedenfalls keine Einheit, kein Segment" (Günther 1988: 76).

95 Note that Neef (2005) chooses the direction letter-to-sound but regards both directions as relevant: "In the end, it is evident that a theory of writing systems has to model regularities in both directions" (Neef 2015: 713).

instance, the doubling of <f>, in a dependent view, signifies the phonological quantity of the preceding vowel and the ambisyllabicity of represented /f/ (cf. Primus 2010: 20–25). In an autonomous, non-linear view (see Primus' model above), the doubling is determined by the structure of the graphematic foot and indicates that the preceding vowel grapheme <ö> is associated with only one (instead of two) skeletal positions. Accordingly, it is the structure of graphematic feet and syllables in German that conditions the variant <ff>. In both views, thus, <ff> is not an idiosyncratic but an explainable spelling (cf. Schmidt 2018: 33; Fuhrhop & Peters 2013: 229–238). In any case, there is no need to assume <ff> as an independent unit let alone an allograph of <f>, which underlines the grave conceptual problems faced by the referential view. However, for <v> and <ph> as written variants of /f/, the picture is not so straightforward. Their status will be discussed in the context of (non-)allography (cf. Section 2.3.4).

The second influential conception of the grapheme in German grapholinguistics is the analogical view. Adherents of it treat the grapheme as the smallest distinctive unit of writing (cf. Fuhrhop & Peters 2013: 202; Günther 1988: 77; cf. Rogers 2005: 10 for an Anglo-American instance of this view). In this view, the phoneme does not serve as the unit the grapheme relates to but instead as a methodological model of how to discover graphemes, which are essentially conceived of as lexically contrastive units of writing. Thus, while the analogical view is not structurally dependent on phonology, it is dependent on it methodologically (cf. Berg 2019: 26). In (earlier) structuralism, the phonemes of a language's phoneme inventory were identified with the help of the substitution of segments and the ensuing discovery of minimal pairs. The same, analogists argue, can be done in writing. Minimal pairs such as German <danken> 'to thank' and <tanken> 'to refuel' reveal that the contrasting units, if they render a difference in meaning and create two existing and meaningful words of a language, are graphemes – in this case <d> and <t>. Thus, graphemes are parallel to phonemes (and morphemes in morphographic systems, cf. Rogers 2005: 10). This approach, Eisenberg (2006: 302) argues, is first and foremost a methodological postulate. This method of "discovering" the grapheme inventory of a writing system is largely uncontroversial if the writing system in question is, ironically, (phonographically) segmental, i.e. a writing system in which a written unit corresponds with a phonological segment. Notably, even in these cases, like the referential view, the analogical view comes with problems of its own.

Most importantly, Kohrt (1986: 88f.) seriously calls into question "whether it constitutes a reasonable project to transfer discovery procedures that have been designed for a specific substantial domain to a totally different one, imputing that in both areas the problems are just the same".[96] Units of speech and

96 Günther (1988: 78, my translation) wonders whether it might be the other way around: "Not the presence of segments is a problem for graphematics, but the lack of segments is a problem for the evaluation of a phonology!" („Nicht das Vorliegen von Segmenten ist ein Problem für eine Graphematik, sondern das Fehlen von Segmenten ein Problem für die Bewertung einer Phonologie!").

units of writing differ in crucial respects, of which Kohrt (1986: 89) highlights *segmentation* and *individuation*. The former proves a much easier task for writing than for speech, as it is one of writing's constitutive features that it is made up of discrete segmental units.[97] This, however, does not mean that all the segments that can be identified as segments are automatically graphemes, as per definition, only units that are lexically distinctive are considered graphemes. Thus, it is individuation that separates graphemes from allographs; it will be discussed in Section 2.3.

It is worth mentioning that Lindqvist (2001: 10) correctly calls out the analogists' double standard of working with minimal pairs that are not truly minimal. Indeed, if minimal oppositions were of interest, the distinguishing contrast between <backen> 'to bake' and <packen> 'to pack' would be subsegmental: what distinguishes them is the spatial (vertical) position of the hasta/head of the basic shapes |b| and |p| within the segmental space. Such subsegmental contrasts are usually not of interest for adherents of the analogical view, who focus on discrete segmental units instead (cf. Gallmann 1986: 47). By contrast, larger, i.e. polysegmental contrasts such as <Qualle> 'jellyfish' vs. <Falle> 'trap' are of relevance for the analogical view. The definition of 'minimality' will be reevaluated in the definition of a universal grapheme concept below.

While the method of discovering graphemes through assembling (segmental) written minimal pairs takes into account that writing is a system in its own right, it is logically divorced from any phonemes the graphemes might relate to or correspond with. The very method itself, which does not require any prior knowledge about a language's phoneme inventory, obscures the fact that while graphemes can distinguish meanings independently, this function may be constituted by phonology (cf. Weder 2016: 13). Is the fact that graphemes that are discovered with this method mostly correspond with phonemes not conspicuous? How much sense does it make to assume that the grapheme differentiates meaning independently of linguistic units such as the phoneme and the morpheme? These latter questions are at the core of the definition of the grapheme proposed here (see below).

Even within the analogical view, grapheme inventories compiled by different scholars often differ slightly. For German, some units such as <sch> are still under discussion; there exist minimal pairs that treat it as one unit as in <Schaum> 'foam' and <Baum> 'tree' but also minimal pairs in which only one component, either <s> or <ch>, is substituted: <Masche> 'bow' vs. <manche> 'some', <Masche> vs. <Maske> 'mask' (cf. Fuhrhop & Peters 2013: 205). In my conception (see below) as well as several autonomous works, <s> and <ch> are clearly treated as two graphemes.

It is paramount to underline that the analogical view does not negate the referential view but rather represents an analysis at a hierarchically deeper

97 As noted in the chapter on graphetics and again below, this often does not hold for handwriting, where the basic shapes are connected to each other, as well as for some scripts such as Arabic, where basic shapes are mostly connected.

level. In other words, it is a logical precursor of the referential view. Evidence for this comes from the assumption of grapheme-phoneme correspondences in the context of the analogical view. Graphemes are gathered analogously to phonemes, but then (and only *then*) are the relations between graphemes and phonemes explored. The correspondences that result from this exploration are the link to the referential view, as they are precisely what the referential view considers to be the minimal units – graphemes. This is in accordance with Günther's (1988: 76) claim that the referential grapheme is not a unit but a relation: it is a phoneme-grapheme correspondence. Interestingly, in Rezec's (2009) model, both the analogous and the referential graphemes are included as separate units: at a lower level, graphemes are assumed in line with the analogous view – e.g. the graphemes <n> and <g> – while at a higher level, so-called *phoneme images* (German *Phonemabbilder*) are postulated as separate units: single graphemes or combinations of graphemes that correspond(s) to a phoneme. An example is <ng>, which, in German, relates to /ŋ/. Rezec's model differs from the traditional analogical view precisely in that the grapheme-phoneme correspondence is treated as a separate *unit* within the model, implying that relations can indeed also be units. The analogical grapheme has also been called *grapho-grapheme* (Heller 1980) or *graphemic grapheme* (Herrick 1994), while the referential grapheme (Rezec's *phoneme image*) has been termed the *phono-grapheme* (Heller 1980) or *phonological-fit grapheme* (Herrick 1994). This shows that there have been several approaches in the past that have attempted to unite both conceptions in one model.

### 2.2.2 Redefining the grapheme

That both the referential and the analogical views cater only to alphabets (and partially other segmental phonographic writing systems) by relying on the phoneme, a unit that is, however, not at the forefront of all types of writing systems, makes necessary an attempt at defining the grapheme cross-grapholinguistically. In advance, however, the central reasons behind this enterprise as well as the possible benefits it entails shall be reiterated.

Firstly, if we take writing as the starting point of analysis – and it is obvious that a theory of writing should do that – it is a simple fact that in every writing system, regardless of how much it differs from other systems, there exists a minimal unit of writing. This does not mean that uncovering this unit or deciding on what it is in each given system are trivial tasks – by all means, they are not. We can start with the fact that all writing at its core functions to encode language. The main question underlying a definition of the grapheme, then, is how minimal units of various writing systems serve this function.

Secondly, from a general theoretical standpoint, more specifically from a linguistic perspective, positing a unit *grapheme* only for a single type of writing system – the alphabet (which Glück 2016b: 251 does, for example) – is a "restriction that would be hard to imagine in the domain of phonology" (Birk 2013, my translation),[98] and this applies not only to phonology but also to many other linguistic domains. Accordingly, it would be untenable to suggest definitions of the phoneme or the morpheme that apply only to one linguistic type or even to only one language. It should similarly count as untenable to proceed this way in the definition of the grapheme. And yet, it appears to be standard procedure in grapholinguistics. Even Kohrt (1986: 91f.), who has produced important work with respect to a grapheme definition, states that he does not believe "that 'logograms', 'morphograms' and 'phonograms' are to be considered as different appearances of one and the same kind and that they should be subsumed under the notion 'grapheme'". Others have not dismissed the idea of a more inclusive and general definition of the grapheme, including Sproat (2000: 25, emphasis in original), who writes "I will use the term *grapheme* to denote a basic symbol of a writing system; [...] I [...] merely use the term *grapheme* as a convenient short way of saying 'basic symbol of a writing system'". Without going into detail about what *grapheme* means for Sproat, it is still noteworthy that he chooses to treat it as applicable to any writing system.[99] I intend to do the same, but given the theoretical orientation of this study, it is necessary to deal with the details. What is the grapheme when we consider vastly different writing systems in its definition?

Three criteria must be met for a unit of writing to be identified as a grapheme: (1) *It must distinguish meaning*. This criterion corresponds with what analogists envisioned the grapheme to be: a contrastive unit of writing. Note, however, that here, it is only one of three criteria that must be met *simultaneously*. Thus, the analogical view, as stated above, was only partially accurate. What is also noteworthy here is that the function of distinguishing meaning does not preclude graphemes to *bear* meaning, which they do in morphographic writing systems in which they relate to morphemes. Trivially, a unit that bears meaning (the morpheme) *automatically* also differentiates meaning,[100] which is the decisive criterion. (2) *It must relate to a linguistic unit (or linguistic information of some kind,* see

98 „[...] eine Einschränkung, wie sie im Bereich der Phonologie nur schwer vorstellbar ist" (Birk 2013).

99 He adds, however, that "in discussing some writing systems we may use the term grapheme in slightly different ways depending upon how fine-grained an analysis is being assumed" (Sproat 2000: 25). This reflects that even though Sproat intends to use the same term for all writing systems, he is aware of the (idiosyncratic) differences it entails in different writing systems.

100 An interesting question that I cannot answer but still want to raise here is whether synonyms that, from a denotative point of view, are semantically identical (if there even is such a thing) could be seen as allomorphs, i.e. different variants of the same underlying morpheme. Allomorphy is most often discussed in the context of grammatical morphemes, but can lexical morphemes with the same meaning but different substances be allomorphs, too? This is relevant for Chinese, because it raises the question of whether synonyms in Chinese that mean "exactly the same" but have different phonological representations could be either written with one grapheme (this would speak for synonyms being allomorphs) or with two graphemes (this would speak for them being separate morphemes). In short, if there are two distinct graphemes and switching them does *not* alter the meaning of an utterance because they relate to mor-

below). This is where it gets a bit tricky and possibly controversial. Of course, in principle, this corresponds with what was claimed by proponents of the referential view. There are, however, notable deviations from the referential view: here, the direction of analysis is, exclusively, from writing to language. How phoneme or morpheme inventories are written, thus, is not the main concern, and thus it is not relevant to identify the written units that correspond to *all* phonemes or *all* morphemes of these inventories. This is also connected to the third and final criterion: (3) *It must be a minimal unit*. Thus, German <ng>, just because it relates to the phoneme /ŋ/ (and thus fulfills the second criterion mentioned here), is not a grapheme because it is not a minimal unit – <n> and <g> are already graphemes. This differentiates <ng> from <ch>: in <ch>, only <h> is an independent grapheme, |c| is not. It is an interesting observation that in the world's writing systems, units that are contrastive but do themselves not relate to linguistic units – such as |c| in German – appear to be rare. Arguably, the definition of the grapheme should not be based on exceptions (but must still be able to explain them), and German, English, and other alphabets have upheld their unjustified status as bases for theoretical models for too long. Table 3 at the end of this section offers examples from various writing systems that highlight how the three criteria apply. First, however, they shall be discussed in more detail.

Criterion (1): Lexical distinctiveness criterion. Graphemes differentiate meaning. Following the analogical view, this condition can be tested using minimal pairs. Notably, both words that are part of a minimal pair must be existing words of a language, i.e. non-words or pseudowords do not count (but cf. Lindqvist 2001). Take German, where this leads to minimal pairs such as <Saum> and <Baum> but also <Schaum> and <Baum>: the contrast between <s> and <b> and corresponding /s/ and /b/ (correspondences that, however, do not yet concern us at this point of analysis) is segmental, whereas the contrast between <sch> (for /ʃ/) and <b> – at least for the first part of the minimal pair – is polysegmental. Cases such as <Schaum> will thus need to be tested by the minimality criterion (see (3) below).

Criterion (2): Linguistic value criterion. While criterion (1) basically corresponds with the analogists' view of the grapheme, criterion (2) partially conforms to what referentialists believe about the grapheme: it corresponds with a linguistic unit, although it is paramount to highlight that this correspondence is not limited to phonemes. What is vital is that while to count as a grapheme, a basic shape must relate to at least one linguistic unit, inversely, not every linguistic unit (such as phonemes in a phoneme inventory) must be 'represented' by a basic shape. Indeed, the analytical direction relevant here is *basic shape → linguistic unit*. This conforms with the analogists' critique that assuming (German) graphemes such as <a>, <ah>, <aa>, and more for the phoneme /a/ or a polysegmental grapheme <ng> for the phoneme /ŋ/ is superfluous.

phemes that are synonymous, does that mean the statement "a unit that bears meaning automatically distinguishes meaning" is false?

Notably, in the graphemes that were discovered using the analogical method, which are expectedly much fewer in number than the referential graphemes, basic shapes also relate to linguistic units; however, this relation was not the defining criterion. Take German again, where the basic shape |c| is, on its own, not the signans of a grapheme. It does differentiate meaning and thus meets criterion (1), cf. <denken> 'to think' vs. <decken> 'to cover' (cf. Rezec 2013: 231). Interestingly, however, even analogists do not claim that |c| is a grapheme; they justify this with its distribution: in native[101] German words, |c| never occurs without <k> or <h>. A possible and rather probable explanation for this distribution that is somehow never directly explicated is that |c| alone (again, in native words) does not represent a phoneme, whereas <ch> /x/ and <ck> /k/ do. While the second of those is not a grapheme because it can be explained suprasegmentally (being syllabically determined much like <ff>, which was discussed above), <ch> *is* a grapheme of German. It differentiates meaning and has a linguistic value by relating to a linguistic unit. But is it minimal? Why is <ch> a grapheme of German and <ng> is not? This is where criterion (3) comes into play (see below).

Criteria (1) and (2) both get fuzzy when it comes to punctuation, which is why criterion (2) does not merely refer to linguistic units but more generally to linguistic value or information. Punctuation marks such as the period or the comma definitely have the potential to distinguish meaning; however, what they relate to remains rather abstract since it pertains to higher linguistic levels – syntax, prosody, pragmatics, among others (cf. Section 2.6). Consequently, criterion (2) was not yet final. A more detailed discussion (that I will not enter into here) will have to reflect critically on whether punctuation marks are to be classified as graphemes and if so, what consequences this has for a grapheme definition. At this point, it might as well be noted that it is not only punctuation marks but other types of graphetic variation that have graphematic potential as well, be it bold or cursive print, underlining, color, or other types of visual design choices. In fact, they have been called graphemes in the past (cf. Gallmann 1985) and were granted, by analogy with prosodic features in phonology, the status of suprasegmentals (cf. Günther 1988). It is safe to say that these latter types of graphematic functions are not encompassed by the grapheme as it is defined here; by contrast, segments such as punctuation marks, digits, special characters, etc. will have to be accounted for. Depending on what exactly a grapheme corresponds with or relates to – whether it is a single, concrete linguistic unit or less palpable linguistic information or a linguistic function – it will be imperative to assume different classes of graphemes that should not be lumped together. I leave this open for future discussion.

Criterion (3): Minimality criterion. To define this criterion, we must first settle on an understanding of 'minimality'. In German, a sequence of two basic shapes such as |ng| is not a grapheme <ng> because myriad minimal pairs can be found in which only one of the segments is substituted: <Bank> 'bank' vs. <bang> 'anxious', <bang> vs. <Band> 'tape', <bang> vs. <Balg> 'brat'. Thus, the fact that

101 It does occur alone in loanwords such as <Clown> 'clown' or <Cello> 'cello'.

the sequence <ng> corresponds with the phoneme /ŋ/ is not decisive since criterion (1) is already met for its constituents (cf. Berg 2019: 30, who argues similarly). In this vein, following Fuhrhop & Peters (2013: 205), the much-debated sequence <sch> should not be considered a grapheme in German since minimal pairs can be found for both <s> and <ch> (see above). The situation is different for <ch> and <qu>. No minimal pairs can be found for their individual components – not for *both* of them, that is. For <qu>, |q| can be substituted in a very limited range of minimal pairs, including <Quelle> 'source' vs. <Duelle> 'duels'. However, the <u> can never be substituted. Thus, |q| alone is not a grapheme because, by itself, it does not differentiate meaning. It also does not meet criterion (2) as it alone does not correspond with any linguistic unit. In such cases where one segment is not a grapheme itself or multiple segments which are combined would not be graphemes individually, units consisting of multiple basic shapes can conceptually still be treated as graphemes; more specifically, they are *complex graphemes* as opposed to *simple graphemes* that consist of only one basic shape. To summarize, what counts as minimal and thus as a grapheme is either a basic shape for which criteria (1) and (2) apply or a combination of basic shapes, in which criteria (1) and/or (2) do not apply for all basic shapes. Combinations such as German <ch> are not graphetically minimal but graphematically minimal.

The above-mentioned problem of subsegmentality can now also be solved. To restate it: if the grapheme is supposed to be the "minimal" contrastive unit, then why are not the head/hasta in |b| vs. |p| or the lowest horizontal stroke of the |E| in |E| vs. |F| graphemes?[102] The answer is simple: neither the hasta (or its location) in |b| and |p| nor the stroke in |E| correspond with linguistic units, meaning they do not meet criterion (2).

**102** In his autonomous, distributional approach, Berg (2019: 27) solves this problem of subsegmentality for the German and English writing systems by preliminarily positing that graphemes must be at least of "the size of a letter". This criterion is problematic since it presupposes (and incorporates) in the grapheme definition a definition of "letter". Thus, graphematic or at least conventional knowledge is necessary in order to know how "large" a grapheme can be. Arguably, in truly autonomous fashion, this problem could be solved by resorting to a purely graphetic criterion, i.e. the spatial cartography of the writing surface. In this case, "at least the size of a letter" becomes "at least a segmental space", whereby the necessity of defining "letter" is avoided and the definition becomes at once applicable cross-grapholinguistically. However, as we will see below (for Arabic and Thai, for instance), this definition is, unfortunately, not accurate, as subsegmental basic shapes can also serve as graphemes. In any case, Berg himself refined his definition of the grapheme and rid it of the problematic "size of a letter" part by relying on the fact that only syllabically autonomous units can be graphemes.

### 2.2.3 The grapheme and subsegmental graphematic components in Chinese

As a morphographic writing system, Chinese is often contrasted with alphabets and positioned at the other extreme of the typological spectrum. Unsurprisingly, thus, its basic units are seldom compared with the basic units of phonographic writing systems. However, as the proposed definition of *grapheme* should be applicable to every writing system, the three criteria suggested above need to be tested for Chinese as well.

Consider the two 'characters' <請> *qǐng* 'please, to ask' and <情> *qíng* 'emotion'.[103] They are complex in that they are constituted by two components: a semantic component (or *radical*), which, in this case, is positioned on the left, and a phonological component (or *phonetic*), located on the right (Myers 2019: Chapter 2). They share the same phonological component, which indicates the pronunciation (except for tone) as it is derived from <青> *qīng* 'green/blue'. Their meaning is differentiated by the differing semantic components, which, therefore, meet the criterion of lexical distinctiveness.[104] As the complexly structured <請> and <情> take up only one segmental space each, according to the analogical view, subsegmental components – both semantic and phonological (for which there are also minimal pairs, see below) – would have to be accepted as graphemes. This is indeed what both DeFrancis (1989) and Sproat (2000) suggest. It constitutes the crucial difference between a polysegmental sequence of two graphemes in an alphabet such as <ng> and the graphematically segmental <請> *qǐng* 'please, to ask': whereas in <ng>, both <n> and <g> retain their status as graphetic segments and individually correspond with a linguistic unit (a phoneme), in the Chinese example, two characters which are graphemes individually – <言> *yán* 'speech' and <青> *qīng* 'green/blue' – are 'crammed' together into one segmental space. The question is, now, whether they still individually meet the linguistic value criterion when they, as minimized versions of the original graphemes they derive from, are part of this new, complex structure.

The semantic components in these examples indicate meaning in an abstract way by signaling an approximate semantic clue. The phonological components indicate the pronunciation of a character (in most cases also only approximately). Prototypical Chinese graphemes relate to morphemes, and neither of the subsegmental components does that: in <請> *qǐng*, the <言> *yán*-part, i.e. the speech-radical, relates to one facet of the morpheme's meaning ('speech'), but it is not in every complex grapheme that semantic components point straightforwardly to the right meaning of the morpheme, as they can also be opaque in this respect (cf. Ho, Ng, & Ng 2003). Examples are <河> *hé* 'river' and <汗> *hàn* 'sweat'. In this

103 The provided pronunciations are from the Mandarin variety of Chinese.

104 However, these components only form minimal pairs with other subsegmental components. <請> *qǐng* 'please, to ask' and <情> *qíng* 'emotion' are a minimal pair, but <請> *qǐng* 'please, to ask' vs. <言> *yán* 'speech' are not.

pair, the water-radical <氵> only indicates semantic affiliation; however, the specific signified of the grapheme <水> *shuǐ* 'water' from which the radical derives is blurred in the process.[105]

The other component in <請> *qǐng*, <青> *qīng*, is a phonological component, i.e. used only for its phonological representation. It contributes a part of the form, i.e. the signans, of the morpheme it originally related to in its full form. For readers, recognizing this morpheme is necessary to access the phonological representation of the phonological component (and thus the clue to pronunciation it provides) even though the meaning of that morpheme is completely disregarded in the final complex grapheme: there is no semantic trace of 'green/blue' in <請> *qǐng* 'please, to ask'. In short, neither graphemes nor subsegmental components of graphemes in Chinese are directly phonographic.

Like <ng>, which conforms to the linguistic value criterion by relating to the phoneme /ŋ/, the graphemes <請> *qǐng* and <河> *hé* correspond with morphemes on their own. However, their subsegmental components, in their minimized sizes, cannot stand alone in the same way that <n> and <g> can. It might be tempting to argue that in a grapheme such as <請> *qǐng* 'please, to ask', the semantic component represents the signatum of the morpheme, while the phonological component serves as the signans of the morpheme. However, as illustrated above, it is not as simple as that: semantic components often only give a vague clue about the meaning of a morpheme, and due to language change, the clues to the pronunciation offered by phonological components have also frequently become opaque (cf. Qiu 2000: 20–21, 247–252; Sampson 1994). At this point, only about 19% of phonological components accurately predict the phonological representation of the morpheme represented by the grapheme (cf. Ho & Bryant 1997: 279).

In this context, consider also the *kanji*-part of the Japanese writing system: *kanji* are morphographic graphemes, many of which are derived from Chinese graphemes. These loaned graphemes often include phonological components. However, since *kanji* correspond with native Japanese morphemes, and these morphemes exhibit Japanese phonological representations, there is no link between originally Chinese phonological components and the native Japanese pronunciations of morphemes (cf. Sampson 1994: 128). These *kanji* are devoid of any phonography, i.e. they are purely morphographic graphemes.[106]

In sum, characters of the type <請> *qǐng* and <河> *hé* are complex graphemes rather than sequences of two graphemes because the two graphemes that are initially independent but are then shrunk in size to fit together into one

**105** In this example, it is also evident that the initial basic shape, if used as a subsegmental component, significantly changes its form. This visualizes the shift in identity from an independent segmental grapheme to a subsegmental part of a grapheme.

**106** Of course, as morphographic graphemes, they correspond with morphemes, and morphemes have phonological representations. These phonological representations, however, are in no way graphematically marked in purely morphographic graphemes.

segmental space (and in this minimized size, cannot occur alone) change their function and cease to meet the linguistic value criterion: they no longer relate to morphemes, and this distinguishes the subsegmental components fundamentally from their individual segmental counterparts, which are undoubtedly graphemes.[107]

What these examples highlight is that there exist features of writing that cannot be explained by phonology or other linguistic levels: the inner systematics of the type of complex graphemes in Chinese characterized above – with one component signaling meaning, the other pronunciation[108] – is a genuine graphematic feature that has no parallel in any other linguistic level. Writing represents language, but that does not mean that all of writing's features should or even can be explained by this representational function.

### 2.2.4 The grapheme in abjads, abugidas, and syllabaries

Now that the three criteria of the proposed grapheme definition have been illustrated with examples from an alphabet and a morphographic writing system, this section will deal with non-alphabetic phonographic writing systems and several challenges they pose for a unified concept of grapheme.

As established in the description of the graphetic module (cf. Chapter 1), *spatiality* is the governing principle of writing. However, thus far, writing has been treated predominantly as one-dimensionally linear (most frequently horizontally), a situation that is referred to as *dogma of linearity* (cf. Krämer 2003: 159). It is for this reason that mainstream linguistics has largely excluded any considerations of two-dimensional features of writing (cf. Waller 1991: 354–357). This can result in the misconception that – in horizontal writing systems – graphemes only appear next to each other and must occupy a relatively equal amount of space. The above-mentioned examples from German and Chinese insinuate indeed that only entire basic shapes, i.e. graphetic segments occupying their own segmental space, can function as graphemes and that they do so only holistically, implying that there are no subsegmental graphemes. However, examples from both *abjads*, in which,

**107** Thus, Daniels' (2018: 168) objections that "there would be two entirely different kinds of grapheme in the writing system, the semantic ones and the phonetic ones, and they would only achieve specific reference in combination" and that "this also leaves the residue of non-composed characters to get some sort of separate description" can be relativized. Not every unit that has a graphematic function is necessarily a grapheme. The fact that 'composed' Chinese graphemes (such as <請> *qǐng*) are complex is something we need to be aware of in a graphematic analysis. However, it does not mean that we must treat them differently than non-composed graphemes (such as <言> *yán* and <青> *qīng*). Both of them differentiate meaning, both of them relate to linguistic units (morphemes), and – as elaborated – both of them meet these specific criteria in a minimal manner.

**108** There are other types of complex graphemes as well, such as graphemes that consist of two or more semantic components (for a typology, cf. Qiu 2000: Chapter 6).

by default, only consonant phonemes and long vowel phonemes are graphematically represented, as well as many *abugidas*, in which default graphemes correspond with consonant phonemes (C) and a vowel phoneme (most often /a/) that is inherent to the consonant grapheme (so that <C> corresponds with /Ca/), appear to contradict this claim. Specifically, they demonstrate that two basic shapes that share a segmental space on the horizontal axis can be separated on the vertical axis.

In both abjads and abugidas, there are vowel graphemes that, on the horizontal axis, are graphetically subsegmental: in abjadic[109] Arabic, they are optional and rarely used, in abugidic Thai, they are obligatory. Take Arabic <رَ> /ra/, which is a combination of the consonant grapheme <ر> /r/ and the (in most contexts optional) vowel grapheme <َ> /a/; they share one segmental space horizontally but not vertically (see Figure 16). The writing system of Thai functions similarly: in a vertical sequence such as <ดี> /diː/ 'good', the two graphemes corresponding with the phonemes /d/ and /iː/, respectively, share one horizontal segmental space. Vertically, however, the basic shape | ี | that materializes the vowel grapheme /iː/ is positioned in a separate space on top of the primary consonant grapheme <ด>. Note that | ี | always materializes /iː/ and can be combined systematically with every consonant grapheme, making it a vowel grapheme < ี >. The fact that the vowel graphemes in these two examples do not occupy their own segmental space horizontally reflects that they are dependent, i.e. cannot occur on their own (Rogers, 2005: 11 calls them *bound graphemes*), at least not in post-consonantal position; note that in some writing systems, they have corresponding allographs that occur word-initially and that are independent (see below).

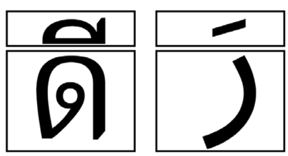


Figure 16. Separate vertical spaces in Thai and Arabic

Aside from the necessary inclusion of the vertical axis, abugidas raise a number of additional questions for a definition of *grapheme*. Their main unit is the so-called *akṣara*, and they are accordingly also referred to as *akṣarik systems* or *akṣara-based systems* (Gnanadesikan 2017: 19). Akṣaras correspond straightforwardly neither

109 Note that in an abjad, by definition, only consonant phonemes and long vowel phonemes are graphematically represented. However, if the optional short vowel graphemes are produced – as in the given examples – both vowel and consonant phonemes are graphematically represented. In this case, Arabic functions more like an alphabet, which renders the discussion of these examples in a section on non-alphabetic writing systems a bit misleading. However, given their size, placement, and dependence, even in vowelized Arabic, vowel graphemes are definitely not equal in status to consonant graphemes. This differentiates vowelized Arabic from alphabets and warrants the discussion of its vowel graphemes together with vowel graphemes of abugidas given that they both share many features.

with phonological syllables nor with phonemes. Instead, they are subsyllabic units whose central component is a long or short vowel that may be preceded but not followed by a consonant or consonant cluster: V, CV, CCV, CCCV, etc. (cf. Patel 2010: 3; Salomon 2007: 28). In phonological syllables, by contrast, vowels can also be followed by consonants. As in the Thai example above, vowel graphemes are commonly smaller in size than consonant graphemes, attached to them (i.e. bound), and are positioned above or below them. However, vowel graphemes can also be of equal size and occupy their own segmental space on the horizontal axis, such as Thai <◌า> which relates to /a:/ and occurs to the right of the consonant grapheme as in <ภาษา> /p$^h$a:sa:/ 'language'. In Tamil, another akṣara-based system, there are also dependent vowel graphemes that are equal in size to the independent consonant graphemes and appear with them linearly on the base line (cf. Bhuvaneshwari & Padakannaya 2014: 192).

As mentioned above, many akṣara-based systems have an additional set of allographs for vowel graphemes that occur initially and are independent of consonant graphemes and equal to them in size: an example is Devanāgarī <उ>, which corresponds with /u/ in initial position as opposed to < ु>, which corresponds with the same vowel phoneme post-consonantally. This raises the question of whether these two variants are two separate graphemes or indeed allographs. Here, the answer appears clear: they are positional graphematic allographs similar to |σ| and |ς| in Greek, which are allographs of the grapheme <σ/ς> and both relate to /s/ in different positions. In short, due to their complementary distribution there exist no minimal pairs between them and they correspond with the same linguistic unit (cf. Section 2.3.2).

In some akṣara-based systems, vowel phonemes are materialized by a discontinuous combination of two basic shapes, i.e. basic shapes that do not appear next to each other but are separated by a consonant grapheme between them; this construction is reminiscent of circumfixes in morphology. Examples come from Tamil as in <ெ·ா> which relates to /o/ or Thai <เ◌ะ> which corresponds with /e/. The question, now, is whether these graphematic representations of vowels are single graphemes or not. In these two examples, the answer is no, and the reason is the above-discussed criterion of minimality. In Tamil, both constituents are already graphemes individually: <ெ> corresponds with /e/ and <ா > corresponds with /ɑ:/.[110] The same is true for the Thai example, where <เ> relates to

110 In this context, an anonymous reviewer (of Meletis 2019) brought up the initial and independent vowel grapheme <ஓ>; just like the discontinuous grapheme combination <ெ·ா>, it corresponds with the phoneme /o/. In this case, given the present conception of *grapheme*, the initial vowel grapheme does not have a dependent post-consonantal *allograph* like other initial vowel graphemes in Tamil do since <ெ·ா> is a combination of two graphemes. However, as this example shows, combinations of graphemes can have non-compositional linguistic values and play important roles in the graphematic representation of the linguistic units (here: phonemes) of a given language. By definition, <ஓ> and <ெ·ா> are no allographs but still graphematic representations of the same vowel phoneme – one being a single grapheme and the other a combination of graphemes.

/e:/ and <ะ> to /a/.[111] Even if these combinations together create a new graphematic value that is not the compositional sum of its parts' functions, they are not complex graphemes but sequences of two separate graphemes, respectively. That way, these examples echo German <sch> that corresponds with /ʃ/ but is not a grapheme since <s> and <ch> are already graphemes themselves. Note that if the two units in such spatially discontinuous combinations were not already both individual graphemes or if only one of them was a grapheme (as in <ch>, where only <h> is a grapheme, rendering <ch> a complex grapheme, see above), such combinations of non-adjacent basic shapes would also constitute single (but complex) graphemes.

Another common feature of these systems is that akṣaras often have consonant clusters in their onsets, i.e. combinations of consonants without a vowel between them. Notably, now, the basic shapes that materialize these conjunct consonant graphemes are frequently contracted to ligatures. In Devanāgarī as used in Hindī, for example, <घ> /gʱ/ and <र> /ɾ/ are contracted to <घ्र> /gʱɾ/. Graphetically, this is reminiscent of the above-mentioned complex graphemes of Chinese in which individual characters that each occupy their own segmental space change their shape (most importantly, become smaller) to fit into one segmental space when they are repurposed as subsegmental semantic or phonological components. The same happens in <घ्र>. The difference between this example and the Chinese examples is that even though the Devanāgarī basic shapes are contracted and are no longer graphetically segmental (and intertwined to such a degree that a visual separation is almost impossible), they still retain the function they had when they were individual shapes – <घ> and <र> –, meaning the ligature still corresponds with two individual linguistic units (phonemes) and satisfies the linguistic value criterion, i.e. it constitutes two (subsegmental) graphemes. The same is the case for irregular, non-standard combinations of consonant and vowel graphemes such as they occur, for example, in Tamil.[112] In sum, these ligatures, which are common in akṣara-based systems, are not single complex graphemes, but sequences of two or more individual graphemes that are graphetically segmental.[113]

**111** <ะ>, however, is also used to graphematically mark the shortness of vowels, which is obviously the function it fulfills in this combination.

**112** Of course, graphetically non-segmentable clusters of graphemes such as the non-standard combination <நு> /nu/ in Tamil are a challenge for the present proposal of a grapheme concept. Instead of analyzing them as sequences of consonant and vowel graphemes that cannot be visually segmented, one could also admit the possibility that aside from consonant graphemes and vowel graphemes, there are additionally also syllabographic graphemes in Tamil (and similar akṣara-based systems) that holistically correspond with phonological syllables. The pros and cons of such an interpretation will need to be discussed in more detail in a specific graphematic analysis of Tamil (or similar systems) that tests the present grapheme definition.

**113** Rogers (2005: 12) calls them *non-structural ligatures* and additionally mentions the typographically motivated ligature <ﬁ> in Roman script, where <f> and <i> also retain their status as graphemes.

Korean – which, despite its 'featural' features (cf. Sampson 2015), is typologically most reasonably characterized as an alphabet[114] – is an extreme example in this respect: here, graphemes, which relate to phonemes, are graphetically subsegmental and are complexly arranged in syllable blocks that are themselves graphetically segmental. The segmental space is thus filled with graphematic clusters that refer to phonological syllables. For example, Korean <ㄱ> corresponds with /k/ and meets all the criteria to be classified as a grapheme; however, it is not used linearly, i.e. is not positioned on the base line and usually does not occupy its own segmental space. Instead, it is combined with other graphemes as in <각>, which corresponds with the syllable /kak/. This graphematic syllable block, although it is graphetically segmental,[115] is no grapheme but a combination of three graphemes. As the subsegmental shapes unambiguously relate to linguistic units (phonemes), it must be noted that graphemes can definitely be graphetically subsegmental as long as they still meet all of the grapheme criteria. Crucially, Korean graphemes differ from the subsegmental graphematic components in Chinese that work only compositionally (see above).

However, it cannot be denied that in Chinese, there is still a relevant graphematic level that is subsegmental, as obviously, semantic and phonological components, elements smaller than graphetic segments can have a graphematic value. Other examples of subsegmental graphematic functions include the |◌̈|-diacritic in umlaut graphemes such as <ö> and <ü> in German that graphematically marks a change of the vowel phoneme represented by the unmarked graphemes <o> and <u>, or the Japanese diacritic | ゛ |, which, when it is added to hiragana graphemes, marks voicing of the initial consonant phoneme in the syllables (or moras) represented by the graphemes: <き> /ki/ becomes <ぎ> /gi/.[116]

Note that this section is incomplete. It is to be expected that writing systems not treated here provide additional open questions for the proposed grapheme definition. These will have to be addressed when they are encountered. However, the fact that the above-mentioned idiosyncratic features of various systems can be explained within the proposed conception points to the probability that the true core features of the grapheme have been identified.

114 Gnanadesikan (2017: 29) calls it a 'fully vowelled syllabically arranged featural segmentary'; this means she regards the syllabic arrangement and the graphemes' partially pictographic representation of places of articulation as features of a phonographically segmental writing system, essentially an alphabet.

115 The segmental syllable blocks of Korean could, of course, also be interpreted as complex arrangements of smaller segmental spaces which are occupied by graphemes. However, due to the many spatial possibilities of combining these graphemes (and the spaces they occupy), it appears more economical theoretically to interpret the syllable blocks as graphetically segmental and the graphemes as subsegmental.

116 These examples imply that there are limits to a segmental graphematic analysis that is based on segmental minimal pairs. As it advances, graphematics (and grapholinguistics in general) will likely evolve in similar ways to phonology and develop more fine-grained featural analyses such as the ones already partially proposed by Primus (2004, 2006), for example. Again, this present proposal is just a starting point.

Table 3. Criteria for the present grapheme definition applied to units from different writing systems; graphemes are highlighted in grey

| | **criterion (1)** | **criterion (2)** | **criterion (3)** |
|---|---|---|---|
| German <ng> | + | + phoneme /ŋ/ | – both <n> and <g> are already graphemes |
| German <ch> | + | + phoneme /ç/ | + |
| German <sch> | + | + | – both <s> and <ch> are already graphemes |
| German \|c\| | + <denkt> vs. <deckt> | – | + |
| German \|I\| (hasta) in <p> vs. <b> | + | – | + |
| Chinese <河> 'river' | + | + morpheme {river} | + |
| Chinese \|氵\| | + | – semantic component 'water', which itself represents no linguistic unit | + |
| Thai <ด> | + | + phoneme /d/ | + |
| Thai <ิ> | + | + phoneme /i/ | + |
| Korean <ㄱ> | + | + phoneme /k/ | + |
| Korean <각> | + | + syllable /kak/ | – is made up of three graphemes that represent phonemes, respectively |
| Japanese <き> | + | + mora /ki/ | + |

The present proposal for a more universal conception of the grapheme is precisely that: a *proposal*, and I do not claim that it is the (only) right one. Like Berg (2019: 32f.), I also want to advocate for more composure in the discussion of grapheme definitions. It does not make sense to speak of "one grapheme inventory" of a given writing system, as different conceptions result in different – and (possibly) equally justified – inventories. Since different analyses are driven by distinct epistemological interests, it would be unfair to proclaim *a priori* that only one method is the right one. All I want to stress is that for a comparison of typologically diverse writing systems, the conception outlined above seems to be a reasonable fit.

## 2.3 Written variation and allography[117]

In language, variation is ubiquitous. Every time a person speaks, variation plays a role at some level, be it in the choice of words or the pronunciation of an utterance. Against this background, it is not surprising that variation has emerged as one of the core phenomena studied by linguistics. Not only speech but also writing is affected by variation. Examples are the difference in the outer form of a written utterance when jotting something down hastily on a shopping list vs. penning something meticulously on a greeting card but also the alternation between uppercase and lowercase graphemes in different positions of a written sentence.

Yet, variation in writing remains a largely understudied topic. More than two decades ago, Coulmas (1996a: 174) observed that there was "no general theoretical model for categorizing graphs as allographs of a grapheme in a given writing system". This situation has scarcely changed. Moreover, what Coulmas mentions is a simplification of the many facets that constitute variation in writing, which is a complex bundle of phenomena. While concepts such as *allophony* and *allomorphy* have entered mainstream linguistics long ago and have developed into clear-defined concepts, the notion of *allography*, much like the grapheme (see above), remains more or less a mystery. Most descriptions of writing systems introduce and work with individual descriptive categories tailored specifically to the writing system in question. This is an understandable reaction to the seemingly insurmountable diversity of the world's writing systems and the associated lack of established and universal descriptive categories in the vein of *phoneme* and *morpheme*. However, settling for individual descriptive categories and altogether avoiding the search for universal methods of describing writing is what created this situation of a conceptual and terminological grapholinguistic vacuum in the first place, and it is only perpetuated when we adhere to this practice. Furthermore, this complicates both the construction of a general theory of writing and, in turn, any reasonable (structural) comparison of diverse writing systems.

The lack of attention that the concept of allography has received concerns primarily descriptive works in which the matter is most often oversimplified. Daniels (2017: 88), for example, after explaining his rejection of the grapheme (see above), remarks that "'[a]llograph, however, remains useful for conditioned variants of lettershapes". However, a definition of 'lettershapes' or an explanation of the mentioned conditions are sorely lacking. Qiu (2000: 297), in his description of the Chinese writing system, writes that "[a]llographs are characters which have the same pronunciation and meaning but have different outward forms. Strictly speaking, only characters which are used in completely the same way, that is, alternate forms of a single graph, can be called allographs". While this is already a more detailed definition, questions do remain: What are *graphs* supposed to be in this context? Are they not commonly understood to be concrete realizations of a grapheme (cf. Section 1.2.1)? In his textbook on writing systems, Rogers (2005: 10–11)

117 This section overlaps largely with Meletis (2020b).

vaguely defines "graphemes as classes of allographs" and adds that "[a] grapheme often has a good deal of allographic variation related to style of handwriting or printing". He obviously speaks of the visual level and thus describes a different phenomenon than Qiu. Roger's brief treatment of allography is surprising given that he claims the "nature of allographic variation and its conditioning factors is more complicated for graphemes than for phonemes". A concept of allography that captures precisely this complexity remains a desideratum.

Aside from descriptive works in grapholinguistics and general linguistics, the lack of a concept of allography also affects psycholinguistic research and specifically questions concerning processes of reading and writing (cf., for example, Rothlein & Rapp 2017). Here, the lack of a definition of different types of allography is detrimental, resulting in a confusing situation in which altogether different phenomena are blanketly labeled as *allography*, which complicates considerably a straightforward comparison of psycholinguistic research on writing and an integration of otherwise valuable findings into more general models of reading and writing. A clear-cut structural definition of allography would allow for a better operationalization in the design of different psycholinguistic experiments and, in the crucial next step, a better interpretation of results and a more theoretically sound assumption of models.

In the following, different types of written variation will be presented that correspond with different types of allography. Crucially, there is a conceptual distinction between variation and allography: Written variation such as in <advisor> vs. <adviser> or <Joghurt> vs. <Jogurt> (German for 'yogurt') might be constituted by differing segments (such as <o> vs. <e>) but the benchmark is the word level. Thus, variation concerns variants at the polysegmental graphematic level rather than the segmental level of graphemes (given that, for instance, <o> and <e> are not always variants, only in specific words). Moreover, it must be noted that written variation can occur in complex sociolinguistic situations and often involves more than one writing system. Examples include traditional vs. simplified Chinese characters that are in use in different systems or also Arabic script as used for Urdu and Devanāgarī as used for Hindi, two languages that are mutually intelligible in speech but separated in their distinct writing. Two versions of the same word written in Arabic (Urdu) or Devanāgarī (Hindi) are not allographs of one another, but variants. In short, *variation* is understood broadly to denote alternations that can be polysegmental and intersystemic (i.e. occur across writing systems). Furthermore, variation may also be determined by (optional) orthographic regulations in a writing system, as is the case for <Joghurt> vs. <Jogurt>. By contrast, *allography* is interpreted narrowly. It captures alternations at the segmental level – either of basic shapes or graphemes – that occur only intrasystemically, i.e. within one writing system. Unlike variation, allography is always constituted by the system itself and is thus never determined externally, i.e. by orthographic regulation.

### 2.3.1 Graphetic variation and allography

Wherever there is writing, there is graphetic variation.[118] Every person who writes by hand has individual handwriting, and typographically, hundreds of thousands of typefaces exist for many of the world's scripts, including Roman, Cyrillic, Arabic, and Chinese.

When writing by hand, each person's handwriting has a specific visual appearance, and so does every typeface. Different people's handwriting and different typefaces (or even just styles of typefaces, e.g. the roman, bold, italic styles of a given typeface such as Times) can be conceptualized as so-called *inventories*. Prior to the writing process, an inventory is fixed either by the fact that the person who is writing by hand has specific handwriting or, in the case of typing, by the choice of a given (style of) typeface. Crucially, inventories are visually distinct: though there may be remarkable similarities between them, the appearance of handwriting inventories usually differs from writer to writer, and typefaces, i.e. typographic inventories, also vary in visual respects, even if sometimes only slightly.[119] Note that even the visual appearance of a *single* person's handwriting inventory commonly varies depending on the communicative writing situation (including formality and the relationship between writer and addressee – in cases in which there is an addressee) but also due to physiological and motor factors such as fatigue (cf. Parush et al. 1998) or external aspects of the writing process such as pen pressure, pen grip, speed, etc. (cf. van Drempt, McCluskey & Lannin 2011).

Since handwriting (or *chirography*) and print (or *typography*) pose different challenges to allography, they are often separated (cf. Fuhrhop & Peters 2013: 207). Once their differences have been dealt with, however, there is no reason not to consider both in the same conceptualization of allography. They differ mainly in the visual similarity between graphs: when writing <cabana> by hand or typing it, six graphs are produced. Obviously, in this word, whether handwritten or printed, three of the six graphs are visually similar or, in print, almost identical because they instantiate the same grapheme (and basic shape). Due to the way it is produced, handwriting is not constant but dynamic in its appearance (cf. Wing 1979). Thus, graphs that are assigned to basic shapes will never look exactly the same. For <cabana>, there is, of course, a theoretical possibility that two or even all three of the graphs instantiating |a| look identical, at least to the human eye, but

118 Here, variation is not only to be understood in a syntagmatic sense in which the different tokens of one type are compared to each other, but also in a paradigmatic sense in which types are related to other types. Even if a person produces only a single graph (for example when writing down only the English indefinite article <a>), this graph would be – in a paradigmatic sense – a variant of all the graphs that the person could have possibly produced to instantiate the basic shape in question. As Spitzmüller (2013: 212) posits, variation is not choosing between different possibilities of graphically communicating 'something'; this 'something' is rather constituted by the choice in the first place.

119 As Hamp (1959: 2) remarks, "[m]any of these [typefaces, D.M.] are characterized in such subtle ways that the average person is not aware of their individuality as such".

the likelihood is much greater that they differ visually at least in some details. This variability of graphs applies more to handwriting than to print, i.e. typographically produced or digitally presented writing, since typefaces are usually visually constant.[120] In print or digital typography, therefore, the possibility of making a conscious style choice at the level of graphs is limited to the choice of typeface itself, and this choice either precedes or follows the writing (or typing) process. Thus, the potential of sociosemiotic self-referencing is not granted directly by the appearance of the typeface, which the producer of a text typically has no influence on, but rather by the choice of a pre-existing typeface.[121] Then, once a typeface has been chosen, the graphs that are materialized and assigned to basic shapes are usually visually constant.

*Graphetic allography* is constituted by the relationship between graphs that are assigned to the same graph classes as well as graph classes that are assigned to the same basic shapes (cf. Figure 17). Visual similarity is a deciding criterion the same way phonological similarity is a deciding criterion in allophony: units must always be visually similar both to each other and the visual configuration of the abstract unit that is located at a higher level in order to be regarded as graphetic allographs. Visual similarity, here, is defined by a similar or identical (1) number of segments that two (dis)similar units are composed of, the (2) relative size of these segments, their (3) arrangement in the segmental space, and their (4) topological configuration, i.e. (dis)connections, junctures, etc. There are two types of *graphetic allography*, and the difference between them is constituted by the notion of *inventory* introduced above.

The first subtype is *intra-inventory graphetic allography*. As mentioned above, the writing process is preceded by the choice of an inventory: when a person writes by hand, the inventory is their handwriting, and when a person writes digitally, the inventory is the (style of) typeface they choose. Against this background, reconsider the production of the written word <cabana>. Whether in handwriting or in print, when this word is materialized, three graphs are produced that are visually (very) similar.[122] Since it is uncommon (although not impossible) to switch to another inventory within the context of a single word (as in ?<cabana>), these three instantiations are members of the same inventory. In this context, a first abstraction can be performed: concrete graphs are subsumed by

120 This claim must be relativized. Even when set in the same typeface, different graphs of a basic shape can differ within the same printed product of writing. Just to give an example: when the ink of a printer is slowly running dry, the color and quality and even shape of the individual graphs on a page might differ noticeably (Andi Gredig p. c.).

121 Note that people simply might not change the default typeface preset in an application, for example Calibri (or in the past Times New Roman) in Microsoft Word. In this case, they have not actively chosen a typeface and might not even have given the use of typeface a thought.

122 Note that all graphs in a (style of a) typeface – not just the ones instantiating the same basic shape – are in a way visually uniform, for example with respect to stroke weight, stroke contrast, and stress angles (cf. Gauthier et al. 2006: 555).

so-called graph classes. *Graph classes* capture the fact that three graphs produced in the same inventory will be visually more similar than |||a||| is to |||a|||, two graphs from Courier New and Arial, respectively. Visually similar graphs within a given inventory are so-called *intra-inventory graphetic allographs*.

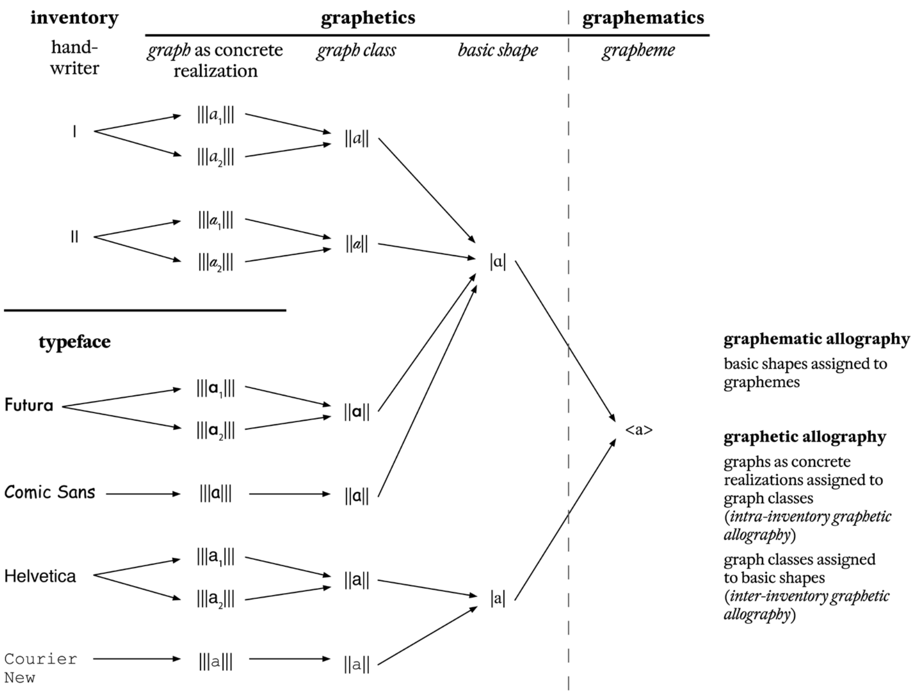



FIGURE 17. Overview of type-token relationships in allography

Intra-inventory graphetic allographs are in a syntagmatic relation with each other; they occur simultaneously on a linear axis in slots in which allographs of the basic shape |a| need to be materialized. Note that they are also in a paradigmatic relation: they are members of the same graph class and, in turn, the same basic shape |a|; however, they are not identical since they are concrete and unique physical shapes. In theory, intra-inventory graphetic allographs are not bound to a given position, meaning they are substitutable for each other within a given syntagma (e.g. the word <cabana> or a larger context such as a document set in one typeface or written in a given person's handwriting). This means they are (relatively) free allographs. However, note that since they are located at the lowest etic level, where, especially in cursive handwriting, coarticulation is of relevance, there are limitations to the notion of 'free'.[123]

123 At the concrete level of production, in handwriting, coarticulation plays a certain role: at least in cursive handwriting, graphs that are produced are connected to each other and may adapt their shape to the graphs that precede and follow. Therefore, even intra-inventory graphetic allographs might not be completely substitutable for

The second type of graphetic allography, *inter-inventory graphetic allography*, concerns allographs from different inventories: in the three instantiations <cabana>, <cabana>, and <cabana>, distinguished by the use of different typefaces, three different graph classes are associated with the basic shape |a|: ||a||, ||a||, and ||a||. Thus, *inter-inventory graphetic allography* is paradigmatic in nature, as ||a||, ||a||, and ||a|| constitute a paradigm, the paradigm 'possible instantiations of the basic shape |a|'. Note that they are not concrete graphs, but graph classes. In larger contexts, e.g. whole layouts in documents, books, etc., inter-inventory graphetic allographs can co-occur if different typefaces are used next to each other, which is common (e.g. sans-serif typefaces for headings, serif typefaces for running text).

Inter-inventory graphetic allography subsumes the above-mentioned intra-inventory allography (cf. Figure 18): for example, in <cabana>, one inter-inventory graphetic allograph (the graph class ||a||) and three intra-inventory graphetic allographs of the basic shape |a| are produced (which are underlined). This is the case regardless of which inventory is used to write the word. The crucial difference between the two types is that, as the name implies, inter-inventory graphetic allographs do not occur in the same context and are determined by the inventory that is used. A Courier New graph |||a||| only occurs in the inventory Courier New, an Arial-|||a||| only in the inventory Arial. To sum up, intra-inventory graphetic allography is largely a syntagmatic phenomenon (and marginally a paradigmatic one), while inter-inventory graphetic allography is exclusively paradigmatic. As they are subtypes of graphetic allography, for both, visual similarity is crucial.

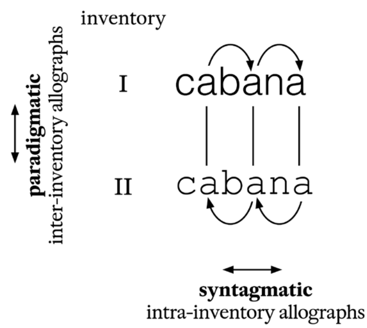



Figure 18. Intra-inventory and inter-inventory graphetic allography

While the term *allography* should be reserved for segmental alternations, graphetic variation occurs not only at the segmental level, i.e. the level of individual graphs. Take as an example the sentence <I do *not* believe this is true.>. Here, the main

each other. In typography, too, there are types of coarticulation such as ligatures: for some combinations of basic shapes, in many typefaces, special connections are programmed, e.g. for the combination of |f| and |i|. Thus, a concrete |||i||| that is produced after an |||f||| might not always be substitutable for a different |||i|||, even if that different |||i||| occurs in the same word as in <finish>.

function of the visual feature *italics*, or more generally, the switch to a different inventory, is to indicate a contrast, to conceptually distinguish the word printed in italics from not only the other words in the sentence but also the other paradigmatic possibilities that could have been produced in its slot, mainly the non-italicized <not>. Its function, thus, is contrastive, and it only works suprasegmentally, since if an italicized word occurred in isolation or all words in a sentence were italicized, no such contrast would be constituted (cf. Meletis 2015: 144–150 and Section 6.8). Even if this contrastive suprasegmental function can be interpreted as linguistic, as it most certainly involves linguistic levels – textual, pragmatic, discourse levels –, it is not regarded as denotative, since the sequence of graphemes materialized by the graphs in <*not*> still corresponds with both the same phonological representation and with the same semantics as the non-italicized <not>.[124] This is one of the central reasons the etic level is so often discarded in grapholinguistic research.

## 2.3.2 Graphematic variation and allography

Whereas graphetic allography is concerned with graphs being assigned to graph classes and basic shapes, graphematic allography is concerned with basic shapes being assigned to graphemes (cf. also Herrick 1975b). Visual similarity was a necessary criterion for graphetic allography, but it is not for graphematic allography. In other words, graphematic allographs can exhibit visual similarity but are not required to. This way, graphematic allographs conceptually resemble allomorphs, which can be but are not necessarily phonologically similar.

To distinguish between several subtypes of graphematic allography, three criteria are relevant:

1. *intra-inventory* vs. *inter-inventory*: As for graphetic allography, this criterion describes whether allographs occur within an inventory or not.
2. *free* vs. *positional*: Positional allographs are complementarily distributed with respect to different positions; their use is conditioned by the system. The (initial) choice of free allographs is free; it is, to a large degree, a stylistic choice.
3. *externally independent* vs. *externally determined*: The default types of graphematic allography are based on graphematics alone, whereas types that are externally determined are determined by other linguistic levels such as syntax or pragmatics. Since externally independent allography is the default, only externally determined types will be explicitly marked in the terminology.

A phenomenon that will not be considered as allographic in the narrow sense is *orthographic variation*, defined as variation that does not stem from the resources of the graphetic and graphematic modules but from the system-external codification of orthographic rules. For the distinction between graphematic and orthographic variation, the additional criterion (4) *systematic* vs. *normative* can be proposed (see

124 Note, however, that the prosodic properties of the entire sentence – if read aloud/ spoken – might be changed.

below). Since systematic variation is the default, only normative variation will be explicitly marked in the terminology.

The first type of graphematic allography is referred to as *inter-inventory free graphematic allography*. In writing systems using Roman script, it is exemplified by the pairs of basic shapes |a| vs. |ɑ| and |g| vs. |ɡ|,[125] respectively. At first glance, these pairs might appear like instances of inter-inventory graphetic allographs. However, crucially, they are not sufficiently similar at the visual level. Basic shapes, including these four, are, as established above, defined by the number of segments they are composed of, the relative size of these segments, their arrangement in space, and their topological configuration. In these regards, |a| vs. |ɑ| differ, as do |g| vs. |ɡ| and |b| as characteristic for print vs. |𝒷| as characteristic for cursive handwriting. By contrast, in order to count as graphetic allographs, two graphs may not differ in the number, relative size, and spatial arrangement of their segments.[126] Basic shapes, thus, cannot be grouped together based on visual criteria. Nothing makes |ɑ| visually more similar to |a| than to |o|; thus, visual criteria could lead to wrong categorizations at the level of basic shapes. This is also what renders this type of variation graphematic rather than graphetic: what is decisive to identify two basic shapes as allographs is that they are assigned to the same grapheme, i.e. that they are functionally equivalent. In most writing systems, |a| and |ɑ| are allographs because they are assigned to the same grapheme <a>. They are not graphemes themselves since they do not differentiate meaning in minimal pairs such as English <ask> and <ɑsk>. This also means they both typically relate to the same linguistic unit in a given writing system, e.g. the phoneme /ɑ/. They can substitute each other, but they typically only do so *across* inventories: they are paradigmatic, i.e. inter-inventory allographs. As such, they are – similarly to inter-inventory graphetic allographs Courier New-||a|| and Arial-||a|| – complementarily distributed with respect to inventories.

Accordingly, when a typeface uses |a|, it will not simultaneously use |ɑ| – with the exception of different styles of typefaces, such as in Times New Roman, for example, where the roman variant uses |a| and the italic variant uses |*a*| (cf. Rezec 2013: 245–247).[127] Accordingly, styles such as bold and italics within one typeface are also conceptualized as distinct inventories.

125 In a recent study, it was found that most participants were unaware of the fact that aside from |ɡ|, there is a second variant, |g| (cf. Wong et al. 2018; Section 7.1.3).

126 There must be some leeway for visual variation, however, otherwise readers would not be able to categorically perceive different graphs as members of the same basic shape. This variation is often of geometric nature, i.e. the relative size/length of segments differs from one graph to another, cf. |A| vs. |A|. In the latter, the horizontal stroke in the middle of the basic shape is shorter in length, which is why the two diagonal strokes are positioned more closely together. This is an example of geometrical variation, which is perceptually less salient than topological variation (cf. Changizi et al. 2006: E119).

127 Cf. Herrick (1974: 11), who states |ɑ| is “considered typical of […] ‘italics’”.

As the designation of this type of allography implies, the *choice* between |a| and |ɑ| is, in principle, free. However, after a choice has been made, the use of either variant is constant. In this vein, Rezec (2013: 245) speaks of consistency rules: this can be illustrated with the example ˀ<egɡ>, which at the very least looks strange to the eye. For handwriting, it seems plausible that people who use |a| stick to it, at least within one text. Their preferred choice of basic shapes can certainly also change as handwriting inventories evolve over time. Also, there exists no explicit (orthographic) rule prescribing that even within one text, a person may not freely alternate between |a| and |ɑ|, although, as mentioned, this is uncommon.

Another well-known example of this type of allography comes from Cyrillic script. Here, the basic shapes associated with a grapheme are conditioned by the use of cursive handwriting vs. print. The default basic shape for the grapheme <т> is |т|, but in cursive handwriting, the visually dissimilar |*т*| is used. Another notable example is |г| and |*г*| for <г>, with the first allograph being used in print and the second in cursive handwriting. Note that this alternation occurs also in typographic writing, as italic styles of Cyrillic typefaces often also use the cursive allographs (as they do here).

In Chinese, there exist so-called *yìtǐzì* (異體字) or *variant characters*, i.e. basic shapes that are in a graphematic relation with the same morpheme.[128] With respect to the classification presented here, they are *inter-inventory free graphematic allographs.* Galambos (2015) provides examples: |峰| and |峯| for *fēng* 'mountain top', |群| and |羣| for *qún* 'group, flock', and |册| and |冊| *cè* for 'booklet' as well as |裏| and |裡| for *lǐ* 'inside'. In these examples, the same two subcomponents of basic shapes are either positioned next to each other, i.e. horizontally, or on top/below each other, i.e. vertically.[129] Even though the different basic shapes consist of the same components, respectively, due to the different spatial arrangement of the components, they still differ visually in a salient way, loosely[130] similar to the way |L| and |T| are distinct basic shapes although they both consist of two straight lines of equal length. Moreover, due to positional constraints, components of Chinese basic shapes commonly change their form when occurring in different positions inside the segmental space (see below). Aside from these examples, which are characterized by the use of the same components, there are other examples such

---

**128** Many thanks to Zev Handel for his helpful answers to my questions about *yìtǐzì,* which have shaped this part of the book.

**129** "[…] the vertically stacked one is viewed as more 'correct' but because it is difficult to squeeze in all the components, the horizontally-arranged one is preferred for readability" (Zev Handel p. c.).

**130** Note that this comparison is reductive since the subcomponents of Chinese graphemes are most often complex, i.e. themselves made up of more simple components (such as lines) and often already have a graphematic function, i.e. signaling meaning or phonological representation, whereas the lines that constitute |L| and |T| are not complex and do not have graphematic functions. Chinese graphemes, thus, are doubly articulated (cf. Ladd 2014: Chapter 5.4.2), whereas the basic shapes of Roman script are not (but cf. Primus 2006 for a different view regarding Roman lowercase basic shapes).

as |詠| and |咏|, two basic shapes that share one component but are distinct with respect to the other; still, they both correspond with the same morpheme, *yǒng* 'to chant'. Evidently, here, different basic shapes are associated with the same grapheme that relates to a specific morpheme.[131] The use of these different variants is not positionally restricted – it is free. However, as is the case for |a| vs. |ɑ|, the use of a variant is conditioned by the inventory that is used, which means it would be strange if two variants of the same grapheme appeared in the same portion of a text. By contrast, it would not be strange in contexts in which more than one typeface is commonly used, e.g. entire layouts.

What is noteworthy at this point is that the choice of variants, even if it is classified as 'free', may not be entirely free but conditioned by sociolinguistic factors. In this context, it is fruitful to consider Bunčić's (2016a) description of so-called *biscriptality*, designating a situation in which two or more variants are being used in one writing system (or across writing systems). From a sociolinguistic perspective, written variation can be diaphasic (conditioned by registers and style), diastratic (conditioned by social strata), diamesic (conditioned by the conceptual dimension of written vs. spoken, cf. Koch & Oesterreicher 1985), diatopic (i.e. geographical), medial (depending on the writing material), ethnic, confessional, etc. Since this section is concerned with the structural dimension of variation, these other conditions for variation are elaborated in the context of the sociocultural fit of writing systems (cf. Section 8.1).

The second type of graphematic allography, *intra-inventory positional graphematic allography*, is the type that is most often mentioned in the literature whenever allography is mentioned. The most popular example of this type is the alternation between the visually dissimilar basic shapes |σ| and |ς| in the writing system of Greek. Since these basic shapes occur in different positions within the same text (that is instantiated in the same inventory), they are of the intra-inventory type. Take the possessive pronoun <σας> /sas/ 'your' in which both shapes occur simultaneously. While both differentiate meaning, there are no minimal pairs |σ| vs. |ς| since they always occupy different positions: |σ| occurs word-initially and word-medially, |ς| only occurs word-finally. Thus, they are complementarily distributed with respect to word position. However, occurring in different positions alone does not suffice to classify these basic shapes as graphematic allographs of one grapheme. For that, it is necessary to establish that both basic shapes are in a graphematic relation with the same linguistic unit: the phoneme /s/. Note, however, that corresponding with the same linguistic unit alone likewise does not suffice to assume allography (cf. Section 2.3.4).

Another well-known example of intra-inventory positional graphematic allography is found in Arabic. Graphemes in the Arabic writing system have up to four positional allographs: there is always a free (or isolated) basic shape of

131 Qiu (2000: 299–300) lists eight categories of this type of allography in Chinese. Note, however, that for most of the examples he gives, the respective allographs cannot be exchanged since only one of them is the "present standard or current [form]".

a grapheme and, since the shapes in Arabic script are always connected, there are connected (or ligated) shapes that are dependent on whether they are positioned initially, medially (in the middle of two other basic shapes), or at the end of a string of basic shapes. For example, the grapheme <ب> has |ب| as its isolated form, |بـ| as its initial form, |ـبـ| as its medial form, and |ـب| as its final form. Note that several graphemes are exceptions: <ا د ذ ر ز و> do not have allographs that connect to the left (which is the writing direction in sinistrograde Arabic script), meaning they only have two allographs, an isolated shape as well as a shape that connects to the right (cf. Majidi 1996: 5).

All examples mentioned thus far concern the alternation of basic shapes, i.e. graphetic units that occupy a whole segmental space of the writing surface (cf. Section 1.2). There also exist, however, instances of intra-inventory positional graphematic allography that are subsegmental. Take the components of Chinese graphemes. Regardless of their function (phonological or semantic), they can change their shape depending on where in the segmental space they are positioned. Radical number 85, the water-radical, usually appears as |水|; however, when positioned on the left of a basic shape, it appears as |氵|. Radicals are not graphemes themselves since they do not meet all the relevant criteria; however, they are subsegmental graphetic units that do have certain graphematic functions (cf. Section 2.2.3).

### 2.3.3 The status of capitalization

Capitalization is a special type of allography that is controversially debated in the literature and concerns only those writing systems whose scripts offer two corresponding sets of uppercase basic shapes (or *majuscules*) and lowercase basic shapes (or *minuscules*). This includes the Roman, Cyrillic, Greek, and Armenian scripts. The central question in this context is whether uppercase and lowercase basic shapes that are conventionally paired together are two separate graphemes or allographs of one grapheme. For English, Sampson (2015: 16, emphasis in original) claims unequivocally – yet *en passant* – that "<g> and <G> would not belong to a single grapheme; [...] the upper versus lower case distinction is significant". For the German writing system, in which capitalization is a more complicated matter, there exist two differing opinions. One of them is that case is lexically distinctive and thus uppercase and lowercase basic shapes belong to two different graphemes. Indeed, minimal pairs can be found: <Arm> 'arm (as in limb) *noun*' vs. <arm> 'poor *adjective*'. Crucially, however, these contrasting words are not paradigmatic since, as instances of different parts of speech, they cannot occur in the same position in a sentence. The second opinion holds that capitalization of words at the beginning of sentences as well as sentence-internally can be explained with recourse to other linguistic levels (cf. Fuhrhop & Peters 2013: 207f.). In any case, there is not one 'capitalization' but rather various types of capitalization in German: the mentioned sentence-initial and sentence-internal capitalization but also capitalization of address pronouns (*Sie/Ihre* '[formal] you, your'), capitalization of proper nouns, capitalization of conventionalized idioms (such as *Schwarzes Brett* 'bulletin board'), and all caps.

Sentence-initial capitalization is what unifies all writing systems whose scripts exhibit a case distinction. In this position, majuscules function to signify the start of a graphematic sentence (cf. Section 2.6) as minuscules cannot be used sentence-initially (cf. Schmidt 2016). Sentence-initial capitalization, thus, is indeed a form of complementary distribution conditioned by position and, thus, an instance of the above-mentioned *intra-inventory positional graphematic allography*.

Sentence-internally and even word-internally, where lowercase basic shapes are the default (cf. Primus 2006: 9), capitalization can also occur, e.g. in the form of all caps. If a whole word in a graphematic sentence is capitalized, as in <I do NOT believe this!>, this represents a form of *suprasegmental graphematic variation*. Like highlighting a string of text in bold print or italics, it changes the visual appearance of a word (or sequence of basic shapes). However, when setting a word in bold or italics, basic shapes are typically kept intact and are still characterized by visual similarity (with some exceptions, cf. |a| vs. |ɑ| in the different styles of a single typeface, see above), which is what makes these types of highlighting forms of suprasegmental graphetic variation. By contrast, changing a string of graphemes to all caps equals substituting lowercase basic shapes for respective uppercase basic shapes associated with the same grapheme. Given that these shapes often do not exhibit visual similarity, this is a graphematic matter.[132]

Capitalization of address pronouns such as <Sie> 'you pl.' can be explained either pragmatically, by arguing that capitalization is an expression of politeness towards the addressee, or simply orthographically, since capitalization of plural address pronouns is an orthographic rule in German. This leaves one case of German capitalization open for discussion: sentence-internal capitalization. Some attribute it to the noun as a part of speech, positing that all nouns require capitalization. However, a more fine-grained syntactic explanation has largely superseded this view: syntactically, heads of noun phrases are capitalized (cf. Maas 1992; Primus 2010: 30).

Evidently, virtually all contexts in which capitalization occurs are determined by external factors, whether syntactic, pragmatic, orthographic, etc. The third of the criteria for graphematic allography listed above, externally determined vs. externally independent, subsumes these different cases of capitalization in German. All of them are instances of *externally determined intra-inventory positional allography*, with the external determinant unspecified since it must be identified distinctly for each type. The alternation between uppercase and lowercase

**132** Note that this may be different in writing systems using Cyrillic script, in which uppercase and lowercase basic shapes for most graphemes exhibit visual similarity since lowercase basic shapes are mostly just smaller variants of uppercase basic shapes, see e.g. |Ж| and |ж| (cf. Lockwood 2001: 309).

basic shapes is deemed positional rather than free since the allographs are complementarily distributed: take a sentence-internal capitalized noun such as <Essen> ‘food’ in <Das Essen schmeckt gut.> ‘The food tastes good’. In this sentence, the lowercase version <essen> (which graphematically represents the verb ‘to eat’) is not orthographically licensed since that spelling would lead to an ungrammatical syntactic representation of the sentence (literally translated as *‘The to eat tastes good’). The random capitalization of individual basic shapes that are not word-initial is also not licensed orthographically, so it would be incorrect to write *<Das EssEn schmeckt gut.>.[133] Obviously, uppercase basic shapes have certain contexts in which they are licensed, whereas lowercase basic shapes – as the default – occur in all other contexts.

To sum up, in the context of this proposal of types of allography, uppercase and lowercase basic shapes do not instantiate distinct graphemes but are allographs of the same graphemes.

### 2.3.4 Non-allographic graphematic variation

In the hierarchy of units of writing systems, allography stops at the level of graphemes: while basic shapes assigned to the same grapheme (and, thus, relating to the same linguistic unit) are allographs, graphemes that relate to the same linguistic unit are *not* allographs. This is a trivial observation: since they are already separate graphemes, they cannot simultaneously be allographs. Yet, they are graphematic variants of some sort. To separate what is graphematic allography from what is not, it is useful to turn to a distinction that Berg (2016a: 17) makes: *graphematic variation in the narrow sense* is variation between written units that is not accompanied by a change in phonological representation, meaning, or categorical structure.[134] Thus, it corresponds with allography as defined here. By contrast, *graphematic variation in the broad sense* means variation between written units that *is* accompanied by changes in phonological representation and/or meaning and/or categorical structure. The pair <far> and <fɑr> showcases graphematic variation in the narrow sense and, thus, allography between <a> and <ɑ>, whereas the difference between <far> and <for> corresponds with a change in phonological representation, meaning, and categorical structure; it is thus an instance of graphematic variation in the broad sense and *not* classified as allography.

Certain misconceptions about allography that are circulating in the literature are closely linked to the referential definition of the grapheme and the associated dependency view. As was established in the previous section, the

133 The all caps version <Das ESSEN schmeckt gut.> is, however, licensed. Here, the uppercase basic shapes cease to have any grammatical function and, as a suprasegmental form of highlighting, serve other functions instead.

134 Categorical structure is the sum of the features of an expression’s morphosyntactic constituent structure, i.e. {noun, singular, ...} for a word like *song* (cf. Berg 2016a: 14).

referential definition is based on analyses of alphabetic writing systems and regards graphemes “as those units that ‘stand for’ phonemes in written language, i.e. as phoneme signs” (Kohrt 1986: 84). According to this view, allographs are those units that are used to write one phoneme, which leads to the theoretically and terminologically dubious situation that allographs are assigned to phonemes instead of to graphemes. Take, for example, the units <f>, <v>, and <ph> in German. They are *not* allographs of a grapheme <f> simply because they can all correspond with the phoneme /f/. <ph> is disqualified on the basis that <p> and <h> are already graphemes individually, cf. the minimal pairs <Hass> ‘hatred’ vs. <Pass> ‘passport’ and <Haar> ‘hair’ vs. <Paar> ‘pair’. Also, when these graphemes occur together word-internally and do not correspond with /f/, they are separated by a morpheme (and syllable) boundary, cf. English <hop.head>, German <Knapp.heit> ‘shortage’, <Desktop.hintergrund> ‘desktop wallpaper’, <Top.hits> ‘top hits’. Also, there exist minimal pairs <ph> vs. <v> that result in a difference in meaning. Crucially, this can be the case both when <ph> and <v> relate to the same phoneme and when they relate to different phonemes: in <Phase> ‘phase’ vs. <Vase> ‘vase’, <ph> corresponds to /f/, <v> to /v/. Even when they both relate to /f/, graphematic[135] minimal pairs are still possible: <Phon> ‘phone’ vs. <von> ‘from’. <f> and <v> can also not be classified as allographs just because <f> always and <v> sometimes corresponds with /f/. Here, too, minimal pairs (even if only few) can be found in which they both relate to /f/, e.g. <Vetter> ‘cousin’ vs. <fetter> ‘fatter (comparative of fat)’, <viel> ‘much’ vs. <fiel> ‘(he) fell’, <Feilchen> ‘little file’ vs. <Veilchen> ‘violet’. Additionally, there are minimal pairs where <v> and <f> relate to /v/ and /f/, respectively, such as <Verse> ‘verses’ vs. <Ferse> ‘heel’ or <Vokal> ‘vowel’ vs. <fokal> ‘focal’. Evidently, <f> and <v> can occur in the same positions in the syllable and the word, although their distributions are not symmetrical.[136]

Another difference between <f> and <v> and the graphematic allographs |σ| and |ς| is that the latter two exclusively relate to the same linguistic unit. This means they do not have the potential to relate to phonemes other than /s/. In the case of <f> and <v>, as established by the examples above, <f> corresponds with /f/, a correspondence that is unambiguous and context-free (cf. Neef 2005: 56), while <v> commonly also corresponds with /v/, a phoneme that is by default represented by the grapheme <w>. Thus, with respect to <v>, because it sometimes relates to /f/, other times to /v/, Neef (2005: 69–71) speaks of an underdetermined correspondence rule. In order to be allographs of one grapheme, |f| and |v| would

**135** Note that this is not technically a minimal pair as there are two basic shapes that together form a contrastive graphematic sequence instead of only one grapheme (such as <f>) that contrasts with <v>. Also, while graphematically, it is the consonant(s) in the onset that form a contrast, phonologically, it is the vowel: <Phon> has the phonological representation /foːn/ while <von> is decoded as /fɔn/.

**136** In word-final position, <v> is very rare. It mostly occurs in the suffix <-iv> as in <attraktiv> ‘attractive’. Note that here, it corresponds with /f/ because of final obstruent devoicing in German. In other forms of the paradigm, it corresponds with /v/ as in <attraktive> since there, a syllable boundary precedes it: <attrakti.ve>.

have to exhibit stable correspondences with the same single linguistic unit, e.g. the phoneme /f/ – just as |a| and |ɑ|, |σ| and |ς|, and |N| and |n| do, respectively. Only a single minimal pair in a writing system suffices to disqualify them as allographs.

For an example from a non-alphabetic, abugidic writing system, take Thai. In Thai, the existence of 42 basic shapes that are in graphematic relations with only 21 consonant phonemes results in a complex multi-grapheme-phoneme correspondence, i.e. a situation in which multiple graphemes correspond with a single phoneme. They are graphemes, however, and not allographs, since there exist minimal pairs in which the words have the same phonological representation but different meanings. Consider <พาย> /phaay/ 'paddle' and <ภาย> /phaay/ 'part (of space or time)' (cf. Brown 1988: 44).[137] The contrast is constituted by <พ> vs. <ภ>. Even though the shapes |พ| vs. |ภ| are in graphematic relations with the same phoneme, they are still part of two distinct graphemes precisely because of the existence of minimal pairs. What must be noted at this point is that an analysis of graphematic variation in Thai is further complicated by the fact that lexical tones are marked in a graphematically suprasegmental manner, i.e. constituted by multiple factors including features of consonant graphemes. Thus, whether two units in Thai are allographs or not can often not be decided simply and solely on a segmental basis (i.e. with minimal pairs) but must be evaluated suprasegmentally (cf. Section 6.6).

### 2.3.5 Orthographic variation

A type of variation that is marginal since it falls out of the realm of graphematics is *orthographic variation*. Here, the fourth criterion of allography comes into play, (4) systematic vs. normative. All types of allography that were described above are systematic; in each case, allographs are licensed units of the graphetic and graphematic modules and the variation between them is an inherent feature of the writing system or its use. Orthographic variants do not fit neatly into this picture – they are normative in the sense of being determined by an external regulation that does not have to correspond to or reflect the inner regularities of the writing system.

In many writing systems, there exist words that have more than one orthographically codified spelling. In some cases, these words not only differ in one segment but are different in more than one respect. Examples are German *<Majonäse> and <Mayonnaise> 'mayonnaise', *<Bravur> and <Bravour> 'bravery', and *<Wandalismus> and <Vandalismus> 'vandalism', where the respective first variants, however, as highlighted by the asterisks, are old variants that were deemed incorrect by the *Council for German Orthography* in 2016 (cf. Duden 2017: 18). Orthographic variants, crucially, are not part of the system if they are decided on by orthographic authorities and are not motivated by users' actual use of the

137 These types of so-called heterographic homophones in Thai are treated extensively in Brown (1988: Chapter 4).

writing system. Note that if two variants are licensed orthographically, users must choose between them and stick to their choice within a text for reasons of consistency. In this sense, orthographic variants are also of the inter-inventory type.

## 2.4 Graphematic syllable and graphotactics

### 2.4.1 German and other alphabets using Roman script

As established in the previous sections, in German grapholinguistics, the paradigm of autonomous graphematics has become predominant. The concept of a graphematic syllable that emerged in this context is demonstrative proof of this. It is paramount that 'syllable', here, is not to be interpreted as 'phonological syllable', but, starting with Primus (2003), as a modality-independent suprasegmental unit in any linguistic modality – i.e. the spoken, signed, or written modality. Specifically, Primus posits respective syllable structures for the spoken and written modalities of German as well as the signed modality of German sign language, which of course is its own language distinct from German.[138] While the general definition of the syllable is founded on an alternation between more and less salient units (cf. Primus 2003: 7) that can be observed for syllabic structures in all modalities, the question of what counts as salient is, by contrast, specific to the modality or its substance: in speech, more sonorous sounds are salient, in sign language, movements are more salient than locations, and in Roman script, shapes exhibiting the feature [+length] are salient. i.e. basic shapes with visible ascenders or descenders that extend beyond the central space (such as |b| or |p|).

This conception of a graphematic syllable dates back to works of the late 1980s (cf. Eisenberg 1989; Naumann 1989) in which it was described that plosive phonemes occurring at syllable edges are represented in writing by basic shapes that exhibit ascenders as in |b| or descenders as in |p|, whereas vowels that appear in syllable nuclei are written with basic shapes such as |a| or |e|. These latter shapes, crucially, occupy only the central space of the line and are thus called 'compact'. Visual length of basic shapes (or parts of basic shapes), thus, indicates phonological syllable boundaries. Independently of Primus (2004, 2006), whose focus is on correlations between graphetic and phonological features, but with strikingly compatible results, Fuhrhop & Buchmann (2009) segment the minuscule basic shapes of Roman script and conclude that all of them consist of a so-

138 Sign languages (such as American Sign Language, ASL) are independent language systems in which the primary (and in fact only) modality is signed, whereas speech and writing are two modalities of one language system, for example English. Correspondences between speech and writing are, thus, system-internal. In contrast, correspondences between sign language and speech or sign language and writing (if what is studied is not a transcription of the signed system, e.g. SignWriting) are system-external, i.e. translations (cf. Meletis 2017; Section 2.1).

called head (or hasta) and a coda. This is actually a finding that goes back to the work of Brekle (1995) and his *hasta+coda-principle*. In |d|, for example, the vertical stroke |l| is the head, while the curve that is attached to it, |c|, is the coda. Partially based on this hierarchical structure of basic shapes, the following criteria are proposed for the definition of the graphematic syllable:

> Every letter has a head (cf. Primus 2004).
> Every grapheme has a coda.
> The head is the vertical segment which spans the central space by the shortest distance and may exceed it.
> The coda is located in only one space (either central, upper, or lower space).
> (Fuhrhop, Buchmann & Berg 2011: 279)

Crucially, the authors interpret the notion of length, which previously had been implicitly conceived of as a binary feature [±length], as scalar rather than absolute (cf. Fuhrhop & Buchmann 2009: 138). Accordingly, all lowercase basic shapes of Roman script can be placed on a continuum of length, the so-called *length hierarchy* (cf. Figure 19). Interestingly, to return to the question of parallels between phonology and graphematics that was raised in previous sections, this hierarchy is claimed to have “the same epistemological status in graphematics as the sonority hierarchy in phonology” (Fuhrhop, Buchmann & Berg 2011: 277).

| Long head | Slant head | Short straight head | | | | Short bent head |
|---|---|---|---|---|---|---|
| | | Connected at the top | Not connected at the top | | | |
| | | Bent coda | Non-bent coda | | Bent coda | |
| b, p, d, g, k h, t, ß, j, f | v, w, x, z, s | m, n | r, l | i | u | a, e, o |
| | | | **Increasing Length** | | | |

Figure 19. Length hierarchy (from Fuhrhop, Buchmann & Berg 2011: 282)

By analogy with the sonority sequencing principle in phonology, a *length sequencing principle* (LSP) is assumed in graphematics: “The graphematic syllable core is occupied by the most compact grapheme. The length of the segments increases monotonously toward both syllable edges” (Fuhrhop, Buchmann & Berg 2011: 283).[139] Examples of graphematic syllables in German that do not violate the LSP are <ver>, <lo>, and <ren> for the word <verloren> ‘lost’ or <le> and <sen> for <lesen> ‘read’ (cf. Schmidt 2014: 254). Notably, graphematic syllables in German and English can also violate the LSP: in German, the basic shapes |s|, |h|, and |y| as well as the combination |tz| violate it, cf. the words <sparen> ‘save’, <fliehst> ‘(you) flee’, <System> ‘system’, and <platzen> ‘burst’, respectively (cf. Fuhrhop &

139 Cf. also Eisenberg (1989: 66), who had already formulated such a principle, although he did not operate with the parameter *Länge* ‘length’, as he called it *Schwere* ‘weight’ instead.

Buchmann 2016: 362–366, 368). In English, |s| as in <speak>, |h| as in <shame>, and |y| as in <rhythm> cause violations. Most of these violations, however, can be explained systematically, and indeed, some of them even appear to have been functionalized for specific graphematic purposes (for details see Fuhrhop & Buchmann 2016; Fuhrhop, Buchmann & Berg 2011; Schmidt 2014).

Historical evidence supports the assumption of a graphematic syllable. As Fuhrhop & Schmidt (2014: 566) illustrate, diachronic changes have led to an emergence and strengthening of the structure that is now characteristic of the graphematic syllable: a demonstrative example of this is the gradual elimination of non-compact basic shapes such as |v|, |j|, and |y| from the syllable nucleus (for more details cf. Fuhrhop & Schmidt 2014). Furthermore, external psycholinguistic evidence also supports the descriptive assumption of a graphematic syllable: specifically, the visual demarcation of syllable structures aids the reading process (cf. the pilot study by Drews 2011 and also Fuhrhop et al. 2016). This is underlined by Eisenberg (2013: 296), who claims that the graphematic syllable serves perception and, thus, the reader.

Up until this point, the graphematic syllable has been presented as an autonomous unit of writing, defined by purely visual means. In the German writing system, it is the smallest unit that can occur on its own.[140] It can simultaneously be a graphematic word and – according to Fuhrhop & Peters (2013: 216) – even a graphematic sentence.[141] Since the graphematic syllable was solely defined with the help of graphetic criteria, the relationship between the graphematic syllable and the phonological syllable as well as the morpheme has not yet been addressed. In this respect, Fuhrhop & Peters (2013: 228) note that in German, on the one hand, there exist many parallels between graphematic syllables and phonological syllables. Frequently, thus, a word will have the same number of graphematic syllables and phonological syllables. Consider German <legen> 'to put, place' (cf. Fuhrhop & Peters 2013: 228). Corresponding with the phonological syllabification in /le.gən/ or /le.gn/ (depending on whether /n/ is interpreted as a consonantal syllable nucleus), a possible graphematic syllabification is <le.gen>, although the morphologically motivated alternative <leg.en> is also possible; both syllabifications do not violate the LSP.

**140** There is one example of a grapheme in German that can stand alone, the seldom occurring vocative-<O> as in <O Tannenbaum!> (which is the original German title of the Christmas song 'O Christmas tree') (cf. Fuhrhop 2008: 199). Nowadays, it can also be written as the sequence <Oh>, cf. <Oh mein Gott!> 'Oh my God!' (cf. Fuhrhop & Peters 2013: 251). Note, however, that even as a single grapheme, it is simultaneously also a graphematic syllable, which means the claim that graphematic syllables are the smallest units that can occur on their own is still accurate.

**141** I argue that in order to count as a graphematic sentence, it would – at least in German, for which the concept was proposed (but not in Thai, for example) – need to also include a punctuation mark. The graphematic sentence will be discussed in detail in Section 2.6.

On the other hand, and this is central in arguing that the graphematic syllable is an autonomous unit of writing, there are noteworthy differences between the graphematic and phonological syllables in German. One of them is found in the marking of vowels: graphematically, vowels are always represented, with the exception of contractions as in <geht’s> *lit.* ‘it is going’ where the apostrophe indicates the omitted initial vowel of the second word, the <e> in <es> of <geht es>. Consequently, the nuclei of graphematic syllables are always vowel graphemes. By contrast, in phonological syllables, sonorants can also serve as syllable nuclei, as seen in the earlier example /le.gn/. Furthermore, in German phonology, null onsets are forbidden, which is why words commonly do not begin with vowels, and glottal stops are almost always inserted before them (cf. Fuhrhop & Peters 2013: 228). Since the glottal stop – which in most conceptions is not interpreted as a phoneme of German (cf. Wiese 1996: 16, 58) – is not represented graphematically, graphematic syllables can have a zero onset, e.g. <alle> ‘all’, <ulkig> ‘peculiar’, <endlos> ‘endless’. Inversely, word-internal syllables in German often exhibit a graphematic onset such as <he> in <Ru.he> ‘silence’, whereas phonologically, the corresponding syllables have a null onset: /ˈʁuːə/.[142] Finally, differences can also be found in the treatment of ambisyllabic consonants such as the /t/ in /ˈmɪtə/ ‘middle’, which is part of both syllables. In writing, the consonant grapheme is doubled as in <Mitte>, and the two instances of <t> are separated by a graphematic syllable boundary between them (cf. Schmidt 2018: 32).

## 2.4.2 Alphabets using scripts other than Roman

So far, the presentation of the graphematic syllable has centered on writing systems that use Roman script (for additional examples from French, cf. Fuhrhop & Buchmann 2016). Thus, the question of what such a syllable-like unit could look like in writing systems that employ other scripts remains open. When considering other alphabets, i.e. segmental phonographic writing systems in which both consonants and vowels are represented, we can observe that in several of them, too, the visual criterion of length plays a similar role as in writing systems employing Roman script. Consider the Greek and Armenian alphabets as well as alphabets using Cyrillic script.[143]

**142** Note that the syllable boundary in <Ru.he> is simultaneously a morpheme boundary (cf. Veronika Mattes p. c.).

**143** The unicase Georgian script is another candidate for such an analysis, as an even greater number of its basic shapes exhibit the feature [±length] than in Armenian. Of the 33 basic shapes, only four are compact: |ა თ ი ო|. Twelve basic shapes have an ascender, |ბ ზ დ ნ პ რ ს შ ჩ ძ ხ ჰ|, thirteen have a descender, |გ ღ ე ვ კ ლ ჟ უ ფ მ ყ ც ჯ|, and four have both, |ტ ქ წ ჭ|. This implies that [±length] could play a role for graphematic syllable structures in the Georgian writing system, although this remains to be (dis-)proven in further studies.

In the Greek script as used for the Modern Greek writing system, [±length] appears to indicate syllable boundaries precisely as it does in German and English (cf. Primus & Wagner 2013: 43). Three minuscule basic shapes of the Greek script exhibit an ascender, |δ θ λ|, eight have a descender, |γ η μ ρ ς φ χ ψ[144]|, and three have both, |β ζ ξ|. Of the remaining basic shapes, six are prototypically used for vowel graphemes, |α ε ι ο υ ω|, while five are used for consonant graphemes: |κ ν π σ τ|. Consider as an example the word <λεξικό> /lɛ.ksi.ˈkɔ/ 'dictionary', syllabified as <λε.ξι.κό>. Here, only the last syllable, <κό> does not straightforwardly conform to the LSP. Note that at this point, the interpretation of the [±length] feature in Greek is only vague, and a further fine-grained (and not purely graphetic) analysis of a possibly existing length hierarchy similar to the one in German might yield that |κ| is longer than |ό|.[145] In other examples, such as <ἑλληνικός> /ɛ.li.ni.ˈkɔs/ 'Greek (adj. masc.)', syllabified as <ἑλ.λη.νι.κός> or <γράμμα> /ˈɣɾa.ma/ 'letter', syllabified as <γράμ.μα>, the same tendency can be witnessed: visual length indicates syllable boundaries. However, determining whether the graphematic syllable can be assumed for Greek the same way it was postulated for German and English is beyond the scope of this book; these few examples merely point to the possibility.

In Cyrillic, admittedly, the picture is drastically different. Considering the basic shapes that are in use in the writing system of Russian, only one basic shape has an ascender, |б|, five have a descender, |д р у ц щ|,[146] and one has both, |ф|. All the remaining basic shapes, regardless of whether they are used for manifesting vowel or consonant graphemes, are compact: |а в г е ж з и й к л м н о п с т х ч ш ы э ю я|.[147] Take the following example:

> <Алфавиты на основе кириллицы являются или являлись системой письменности для 108 естественных языков.>
>
> 'Alphabets on the basis of the Cyrillic script are or have been writing systems for 108 natural languages.'

---

**144** In some typefaces (such as the one used here), |ψ| and |φ| can have an ascender, too. Prototypically, however, these basic shapes only have a descender.

**145** In Greek, as is evident from these examples, stress is graphematically marked with a diacritic. This adds (albeit arguably non-salient) visual information in the linear space's high space, which, possibly, could also affect an analysis of the feature of [±length] (or similar features).

**146** It is a legitimate question whether the 'hooks' in |д ц щ| are in fact descenders given that they do not extend into the lower space of the line as much as other descenders (such as those in |р у|) do (cf. Gordon 2006: 301). Therefore, Gordon (2006: 36) prefers the term 'dangling' instead of 'descending' for them. Notably, for |д|, he also treats the hooks as 'vertical serifs', which is interesting as for Roman script, serifs were classified as non-distinctive (cf. Primus 2006: 9). If this applied to the Cyrillic script, too, then a |д| without the hooks would only be a (stylistic) 'sans serif' variant. For |щ|, however, interpreting the hook as a serif does not work, as eliminating it results in a distinct basic shape that is used for a different grapheme: |ш|.

**147** Like the marking of stress in Greek, the diacritic on |й| may be interpreted as an expression of [+length].

When visually analyzing this Russian sentence, it becomes obvious that the units that are separated by empty spaces of the second order ('word spaces') lack any visually salient internal organization. In <алфавит> *alfavit* 'alphabet', |ф| does (in this case even doubly) exhibit the feature [+length], and here, it correlates with a syllable boundary: <ал.фа.вит>. However, the other long basic shapes do not share this function, especially |у|, which is in a graphematic relation with a vowel (= a syllable nucleus) and greatly disrupts the picture, cf. <русский> *russkij* 'Russian', which is syllabified as <рус.ский>. In this word, |р| does not violate a potential LSP of Cyrillic script, but |у| does so quite ostentatiously. Like the analysis for Greek, these observations are highly preliminary as no sophisticated segmentation and subsequent hierarchization of segments of Cyrillic basic shapes have been undertaken yet. Even if at first glance, visually, it seems improbable, there could hypothetically still be a length continuum in Cyrillic, too. However, even if such a continuum existed, it would not be as visually salient as the continua for Roman and Greek characterized above.

The Armenian script – which has both upper- and lowercase basic shapes – is noteworthy in that of 39 lowercase basic shapes, only seven lack an ascender or descender. These are |ա ո ռ ս տ օ ւ|. Interestingly, only four of them are used for vowel graphemes, while the other three are used for consonant graphemes. Sixteen letters have descenders, |բ գ դ զ ը թ լ ղ յ շ չ պ ջ ր ց ք|, ten have ascenders, |ե է ժ ծ հ ձ ճ մ ն և|, and six have both, |ի խ կ վ փ ֆ|. When analyzing the syllabic structure of Armenian – only roughly, i.e. again without a subsegmental analysis of basic shapes' constituents –, length appears to play some role in graphematic syllabification: for example, in the word <համալսարան> *hamalsaran* 'university', syllabified as <հա.մալ.սա.րան> <ha.mal.sa.ran>, long basic shapes only occur at syllable edges. The same holds for <վերարկու> *verarku* 'coat', syllabified as <վե.րար.կու> <ve.rar.ku>. Here, the second basic shape and simultaneously the nucleus of the first syllable, <ե>, has an ascender. In this case, notably, this does not constitute a grave violation of the LSP since the nucleus in this case appears syllable-finally. If, however, the syllable were to instead feature an additional final basic shape (especially a compact one), this long basic shape that manifests a vowel grapheme would definitely count as a violation of the LSP. When considering additional examples, it becomes clearer that only a few basic shapes appear to violate the LSP, as in <անձնագիր> *ɑnd͡znɑgiɾ* 'passport', syllabified as <անձ.նա.գիր> <ɑnd͡z.nɑ.giɾ>, where it is the basic shapes |ն| and |ի| that violate it. The tendency revealed by this preliminary analysis is that length is of some importance in signifying graphematic syllables in the writing system of Armenian.

Arabic and Thai are interesting with respect to the investigation of a graphematic syllable in that they are both segmentally phonographic but not alphabetic writing systems (cf. Section 2.7). Visually, in these systems, there are no discernable units below the one-dimensional graphetic sequence that is usually functionalized to correspond with words (in Arabic) and syntactic units (in

Thai).[148] Of course, phonological syllables have written correspondences in these systems, but these are not demarcated in a visually salient way.

### 2.4.3 Syllabaries

While in alphabets using the Roman, Greek, Armenian (and possibly Georgian) scripts, the visual feature that distinguishes salient from less salient basic shapes is functionalized to signify syllable boundaries, this reliance on visual salience becomes a problem for the adoption of a more universal perspective. To recap, an alternation of degrees of salience between units is the decisive criterion in a modality-indifferent definition of the syllable. Notably, salience is defined substantially – for writing, that means visually. Consequently, what is salient across diverse writing systems and the different scripts they employ is subjected to vast visual variation – variation that is unprecedented in the spoken and signed modalities of language. This is likely owed to the fact that while speech and sign language are produced only with parts of our bodies – oral articulators, arms, hands, etc., writing is additionally reliant on external tools. Crucially, our articulators, mainly our hands, in combination with those tools, e.g. pens, do not appear to be subject to the same limits as the 'mere' bodily articulators in the production of spoken and sign language. In any case, material variation is much richer in writing.

Admittedly, Fuhrhop & Buchmann (2009: 152, my translation) do underline that their proposal of a graphematic syllable only applies to German and suggest further studies for other writing systems with the motivation of arriving at universals:

> […] we can make statements about the structure of the graphematic syllable in German, the same can be done for other languages. Afterward, these statements are merged and one possibly arrives at universals (especially with writing systems that operate with the Latin alphabet, other alphabets would be the next step, and in turn, the next step would be the comparison with non-alphabets since, in Japanese, the syllables are already the graphemes).[149]

---

148 McCawley (1994: 122), in his treatment of graphotactics, describes a graphic syllable for Hindi. Its definition resembles the graphematic syllable's as proposed by Fuhrhop and her colleagues as it is explicitly divorced from the concept of the phonological syllable: "The use of the term 'graphic syllable', it should be noted, does not carry with it any presupposition that each graphic syllable represents a phonological syllable; in Hindi, for example, it is common for a word to contain more graphic syllables than it has phonological syllables […]".

149 „Wir können […] Aussagen machen über die Struktur der Schreibsilbe im Deutschen, das Gleiche kann mit anderen Sprachen gemacht werden. Hinterher fügt man dies zu Aussagen zusammen und kommt möglicherweise zu Universalien (insbesondere mit Schriftsystemen, die mit dem lateinischen Alphabet operieren, andere Alphabete wären dann der nächste Schritt, der übernächste Schritt der Vergleich mit Nicht-Alphabetschriften, im Japanischen sind die Silben ja schon die Grapheme)."

With their statement that “in Japanese, the syllables are already the graphemes”, the authors are referring to the fact that the basic shapes of the Japanese *kana* inventories (hiragana and katakana) are in graphematic relations with phonological syllables (or actually moras, but that is another story). They are one type of grapheme in Japanese, with the other type being the morphographic *kanji*. At first glance, it appears correct to claim that the graphemes ‘are’ – or better ‘represent’ – syllables. However, in this quote, the terminological vagueness is fatal, as Fuhrhop & Buchmann (2009: 152), when speaking of Japanese “syllables”, presumably do not mean graphematic syllables but phonological syllables. This, then, strays from the fundamental claim that graphematic syllables can be defined *without* recourse to phonology. As I will show in the following, in Japanese, they cannot.

Let us proceed the same way that the authors did in their assumption of a graphematic syllable in German. That is, in a methodologically autonomous manner in which it is not assumed that graphematics is dependent on phonology (cf. Section 2.1). If we investigate a string of basic shapes from the Japanese hiragana inventory merely graphetically, no visual features can be identified that could straightforwardly be classified as rendering basic shapes more visually salient than basic shapes from the other inventories (either kana or kanji). Take, as an example, the sentence <お腹が空きました> in which hiragana basic shapes are underlined and kanji basic shapes are not. When considering only the visual form of the shapes, then – except the fact that the kanji basic shapes are visually more complex than the hiragana basic shapes in that they are composed of a greater number of strokes[150] – no systematic visual difference can be made out between the two classes of basic shapes. In other words: no visually salient “alternation structure” that was claimed to be constitutive of syllables becomes apparent. Thus, there are no different classes of basic shapes in Japanese that showcase visual features similar to [+length] and [–length] which are largely characteristic of basic shapes that materialize German consonant and vowel graphemes, respectively. This could lead to the preliminary assumption that there exists no “graphematic syllable” in Japanese. The same applies to all other writing systems that are typologically non-segmental (cf. Section 2.7). In Japanese, the segmental units of writing, the graphemes, correspond with either syllables or morphemes. The obvious conclusion, in turn, is that the linguistic level that the written units correspond with (phoneme, syllable, morpheme, etc.) determines whether there can possibly be a visual alternation structure. And here, it is only segmental writing systems, i.e. those in which graphemes correspond with phonemes, that can develop graphematic syllable structures.

In the syllabographic part of the Japanese writing system, there is a transparent and uniform relationship between basic shapes and phonological syllables, which renders graphemes (mostly) biunique (cf. Sections 6.4 and 6.5). However, the kana inventories are complemented by the morphographic kanji, leading to Japanese’s unique typologically mixed nature. This is the reason that

150 While this might be true for the *kanji* and *kana* in this example, there are also simpler *kanji* with fewer strokes that resemble the *kana* basic shapes more closely.

there are two types of graphemes in Japanese, with kana graphemes corresponding with syllables and kanji graphemes with morphemes. Only the empty spaces between kana graphemes consistently indicate (phonological) syllable boundaries. By contrast, since kanji graphemes correspond with morphemes, and native *kun* readings of morphemes are (often) polysyllabic, spaces between kanji graphemes do not always correspond with phonological syllable boundaries (cf. Bhide 2015: 2). Thus, unlike Chinese hanzi, Japanese kanji are not (necessarily) morphosyllabic (or more precisely morphosyllabographic) but instead only morphographic.

The question remains of whether the segmental graphemes that correspond with phonological syllables are also 'graphematic syllables' in the sense presented above. This entails a different crucial question: whether graphemes can simultaneously be graphematic syllables. The definition outlined above clearly states that the graphematic syllable is a 'suprasegmental' unit, with this suprasegmentality being defined visually based on an alternation between salient vs. less salient basic shapes. Like in phonology, where a single phoneme can (and commonly does) serve as a phonological syllable nucleus, a single grapheme also can, in some systems, function as a graphematic syllable nucleus, the only obligatory part of a syllable. Consider the grapheme <a> in English, which is the graphematic representation of the indefinite article as in 'a cat'. This grapheme could be simultaneously regarded as a graphematic syllable and a graphematic word.[151] The same goes for the above-mentioned German <O>. There are several such words that consist of only one grapheme across writing systems, such as French <y>, Spanish <e>, or Dutch <u> – all of which, notably, are function words (cf. Fuhrhop & Peters 2013: 251).[152] This implies that segments, both spoken and written, can be syllables and even words if they fulfill the minimal requirement of being a licensed syllable nucleus. However, against the background of Primus' modality-indifferent syllable definition, I object to the idea of segmental graphematic syllables and propose a narrower, materially defined reading of 'suprasegmentality', which I call polysegmentality.

A crucial observation concerning frequency and markedness is that in phonology, segmental phonological syllables – that is, syllables that consist of only one phoneme – are exceptions.[153] The same applies to single-grapheme graphematic syllables in segmental writing systems such as German or English. In the Japanese writing system, however, for the *kana* inventories, but also for parts

151 Arguably, this multiple identity of <a> is not split in equal shares: <a> is primarily a graphematic word, and only then a grapheme, and finally, and least saliently, a graphematic syllable.

152 As Evertz (2016: 393) notes, in English, the so-called *three-letter-rule* states that a content word must contain at least three letters. Examples of this are <egg>, <bee>, and <pea>.

153 This is accurate at least in that segmental syllables are far outnumbered by polysegmental syllables. Few phonemes in the respective languages of the world have the capacity to serve as syllable nuclei; they are, for the most part, vowels (cf. Vennemann 1988: 27–30).

of the kanji inventory, graphematic syllables that are composed of only one grapheme would be the rule. Notably, according to Primus' syllable definition, these cannot actually be graphematic syllables since syllable structures are constituted by an alternation of salience, which can only be attained in syllables consisting of more than a single unit. Thus, perhaps counter-intuitively, one can conclude that there exists no visually discernible polysegmental graphetic unit in Japanese that could be treated as a graphematic syllable similar to the graphematic syllable in German. Indeed, their units' correspondence with phonological syllables – which lends such writing systems the name syllab(ograph)ic writing systems or syllabaries – cannot be decisive in accepting that segmental graphemes could simultaneously be graphematic syllables, as relying on this criterion would deviate from the autonomous, inner-graphematic analysis that characterizes the German graphematic syllable. In a nutshell, the conclusion for syllabaries reads: graphemes correspond with phonological syllables – they are syllabographic graphemes. Strikingly, this means that in syllabaries, there is no graphematic syllable as defined above.

The terminological and conceptual difference between 'suprasegmental' and 'polysegmental' that was introduced above needs to be further explained, as it is important in the rejection of a graphematic syllable in Japanese and other syllabaries. Any sequence of segments, such as three-grapheme English <all>, is polysegmental. 'Polysegmental', thus, refers merely to the fact that a unit consists of more than one segment. Now, the definition of the graphematic syllable for Roman script depends on the visual length of basic shapes. This feature is not suprasegmental (in the traditional sense of the word), however, such as vowel length in phonology (cf. Fox 2000: 12–14), because it cannot be recognized exclusively in relation to other units. A <t>, when analyzed with descriptive graphetic methods (cf. Section 1.2), always exhibits the feature [+length], an <e> always [–length]. Indeed, although Fuhrhop & Buchmann (2009) conceived of these terms as scalar, it is the length hierarchy that is scalar, while the basic shapes' positions on the hierarchy are absolute. Yet, the definition of 'graphematic syllable' that stems from German literature does not commonly operate on single segments, as it is defined as a polysegmental unit that must (largely) conform to the LSP. One-segment graphematic syllables such as English <a> may not violate the LSP, but, as stated above, they are certainly not the rule. If there were only or even mainly one-segment graphematic syllables in the writing systems of German or English, an assumption of the graphematic syllable as defined above would be grossly redundant.

Now, for Japanese, it may be argued that there does exist a salient visual marker that allows discerning graphematic syllables: an empty space between them. This view is backed by the argument that in light of the visual diversity of the world's many scripts, a broader spectrum of visual criteria that signify graphematic syllables should be considered. In other words: 'length' as defined for Roman script cannot be the sole criterion. Nonetheless, what prevents the empty space between basic shapes in Japanese to be interpreted as a graphematic syllable boundary is the above-mentioned polysegmentality. Segments are already graphemes and, *visually* speaking, they cannot be graphematic syllables (characterized by the varying salience of their constituents), too.

This, then, raises another question that is central to theoretical grapholinguistics: if units are defined visually and without recourse to linguistic units, i.e. *graphetically*, which is how the 'graphematic syllable' is defined, then why are they not conceptually and terminologically treated as *graphetic* rather than *graphematic* units? In other words: is it not a graphetic syllable rather than a graphematic syllable? And is the feature [±length] not merely an additional graphetic feature that allows a further segmentation of internal chunks of visual material within graphetic units that have already been identified with the help of the empty space criterion? If all this is affirmed, then the graphetic syllable is a polysegmental graphetic unit *inside* another polysegmental graphetic unit (the one-dimensional graphetic sequence), which itself is demarcated by empty spaces (cf. Section 1.2.2). If, now, in the next analytical step, we were to find that these graphetic syllables correlate predominantly with phonological syllables like the 'graphematic syllable' does in German and English, then is what has been discovered actually an autonomous unit of writing or instead a reflection of language (and some of its specific features) in writing? If the latter scenario is accepted, then what we have gained is actually a strong argument in favor of the view that writing is structurally dependent on language, and in this case, phonology. Notably, this in no way makes the autonomous methods that led to this conclusion obsolete or unreasonable, since some features may only be captured with such a methodology. Accordingly, strong claims such as Zifonun et al.'s (1997: 263) that a 'written syllable' (= *Schreibsilbe* in the German original) can *only* be defined with recourse to the phonological syllable and cannot be identified with graphe(ma)tic features or combinatorial rules must also be rejected. As the previous discussion emphasized, the intra-graphematic (in essence, this means graphetic) analysis is indeed fruitful and tenable. However, in a further analytical step, the overlap with phonology or other linguistic levels cannot and should not be explained away.

The previous remarks also bring up the question of why a graphematic syllable defined by visual salience is not as universal as the phonological syllable or syllabic structures in sign language. Is the reason for this "un-universality" of the graphematic syllable based on substance, i.e. on the fact that only some scripts exhibit the necessary visual resources such as different classes of basic shapes (e.g. distinguished by the feature [±length]) that allow a visualization of structures in which constituents alternate in their degree of salience? If this is excluded as a reason (and I am not claiming it should be), other possible reasons must be considered. All in all, the type-specificity of the graphematic syllable as described above is striking: only alphabets show this kind of visual alternation in the written modality – an alternation that largely corresponds with the alternation of salient vs. less salient units in speech.

## 2.4.4 Graphotactics

Polysegmentality is not only a prerequisite of syllabic structures in writing, it also leads to a necessary examination of the concept of *graphotactics* or *graphotactic constraints*, defined as "restrictions on ways in which the elements of a writing system may combine with each other" (cf. McCawley 1994: 115; cf. also Voeste 2004 and Balestra 2017 for an overview of definitions of graphotactics). Thus, among other things, graphotactics constitutes rules that state which combinations or distributions of basic shapes as well as graphemes are possible (or 'grammatical') in a given writing system (cf. Stalph 1989: 23). At a different level, it evaluates combinations of elementary forms within basic shapes (such as subsegmental components in Chinese graphemes, see below). In other words, graphotactics is necessary at the graphetic level for the formation of well-formed basic shapes as well as at the graphematic level for the formation of well-formed units (or better 'strings') of writing that are larger than the grapheme. These larger strings do not start with the graphematic syllable, however, as *any* sequence of two or more units can be assessed graphotactically, regardless of whether it constitutes a graphematic syllable or not. A central question here is, once again, whether graphotactic constraints echo phonotactic or morphotactic constraints or whether they can be described autonomously (which is claimed by Günther 1988: 77, for example; cf. also Berg 2019: 33). Arguably, as in the definition of the grapheme, both views are partially accurate. Graphotactics has been studied only scarcely, meaning descriptions of the graphotactic constraints of various writing systems prove even sparser than works on other grapholinguistic problems.

Just as there exist two broad types of allography (cf. Section 2.3), there are also two types of graphotactics. This is echoed in McCawley's (1994: 115) study in which he analyzes graphotactics on three different levels: "I will consider here three different kinds of units: letters, punctuation marks, and graphemic constituents of letters". The final level listed broadly deals with the purely visual, i.e. graphetic restrictions underlying the combination of features within basic shapes but also with combinations of individual basic shapes. This type is termed *graphetic graphotactics*. The second type, *graphematic graphotactics*, deals with restrictions on combinations of graphematic units, i.e. restrictions originating from the fact that writing represents language.

As mentioned, *graphetic graphotactics* deals with the combinatory restrictions on visual material. A description of the graphotactic constraints of a writing system needs to start with an approach similar to Bredel's (2011: 19) or Jacobs' (2005). As laid out in Section 1.2.1, Bredel decomposes the writing surface into three subspaces: the segmental space, the linear space, and the areal space. At a hierarchically higher level, the holistic space can be additionally assumed (cf. Meletis 2015: 115); it is composed of various areal spaces such as paragraphs on a page, with the page being the holistic space. With this modification, the writing space is overall conceptualized as fourfold structured. All its spaces are of relevance for graphetic graphotactics, although the segmental and linear spaces are clearly most crucial. What Bredel's (2011: 19–22) approach, which was designed specifically for German, makes possible is determining in which way classes of basic shapes – such as letters (or more generally, scriptual basic shapes), digits, special characters, and punctuation marks – may combine to form larger units. In addition to these visible basic shapes, empty spaces occur, which Bredel treats as yet another class of 'graphic material'.[154] The empty space that occupies its own segmental space in German is used to visually separate graphematic words from each other (cf. the next section). Bredel's focus is on explaining which punctuation marks occupy their own segmental space on the one hand (*fillers*) and which cliticize on a different segment and thus do not occupy their own segmental space on the other (*clitics*; for details, see Section 2.6). All other types of basic shapes besides punctuation marks occupy their own segmental spaces.

What Bredel's approach also illustrates is that most classes of basic shapes cannot combine across classes: digits can combine with each other to form larger units (i.e. numbers), as in <911>, and letters can, of course, be combined with each other to form words as in <house>. By contrast, punctuation marks, with exceptions such as <?!>, and special characters, cannot be concatenated to build larger units, and digits and letters also cannot combine with each other quite so freely (at least in orthographically correct writing). Notably, at this point, nothing has yet been said about *how* different basic shapes – if they are used graphematically – can combine within their classes. Also, the basic shapes' individual graphematic functions were not mentioned. However, the questions of how basic shapes of different classes behave within their classes and outside of their classes are arguably not solely graphetic, since knowledge about which basic shape belongs to which class is necessary, and as I have argued (cf. Section 1.2.1), this is not solely graphetic knowledge.

As scripts are visual systems, specific rules of composition distinguish well-formed basic shapes from ill-formed ones. Crucially, well-formed shapes can either actually exist as units of a script or be only hypothetical basic shapes structured according to the features of the script (cf. Watt 1983a and Section 5.1). While the necessity to create new basic shapes for existing scripts does not arise frequently, this does happen, for example in Chinese (cf. Zhao & Baldauf 2007). In this case, graphetic graphotactic constraints become relevant. In Chinese, they subsume restrictions about where subsegmental components (radicals and phonetics, cf. Section 2.2) can be placed within the segmental space with respect to each other. Shapes that do not conform to graphetic graphotactics are ill-formed.

The second type, *graphematic graphotactics*, evaluates how graphemes may be combined. This corresponds with the common (and underdifferentiated) reading of *graphotactics*. At this level, it is not basic shapes that are of concern but the graphemes they materialize. Whether grapheme combinations are licensed

154 She does state, however, that it is yet unclear what semiotic status empty spaces have, as they are not exactly 'graphic material' given that they are themselves blank (cf. Bredel 2015).

is to a large degree determined by how the linguistic units they correspond with may be combined. Intra-graphematic factors play a role, too: Günther (1988: 77) provides the example that in German, <h> after <i> occurs only in pronouns and argues that this is a graphotactic constraint that is not linked to German phonotactics. This highlights that writing systems exhibit idiosyncrasies not only at the graphetic level but also at the graphematic level, idiosyncrasies that can only be reasonably described and explained "intra-graphematically". Is this the case in every writing system? And what is the reason for or basis of these idiosyncrasies?

What can be observed is that graphotactics and phonotactics (in phonographic writing systems) or graphotactics and morphotactics (in morphographic writing systems) drift apart considerably when graphemes, i.e. graphematic relations between basic shapes and linguistic units, are not transparent and uniform semiotically, which is the case in many writing systems (cf. Sections 6.4 and 6.5). The less transparent or uniform graphematic relations are, the larger the graphematic solution space of a writing system becomes. At this point, not only orthography comes into play, regulating which spellings are regarded as correct, but also graphotactics, as every writing system develops its own sets of combinatory restrictions in order to cope with the one-to-many relationships of basic shapes and linguistic units. The more biunique the relationship between them gets, however, the more redundant an autonomous graphotactics becomes: in Chinese, for example, where there is a nearly biunique relationship between basic shapes and morphemes, there are almost no graphotactic constraints restricting the combination of graphemes that are not determined by the morphotactic constraints affecting the morphemes that the graphemes correspond with.

## 2.5 Graphematic word

The seminal work in defining the graphematic word is Fuhrhop (2008). As in the context of the graphematic syllable, Fuhrhop cautiously states that "the term graphematic word is established here for German. Some aspects will be generalizable, others will not" (Fuhrhop 2008: 190, my translation).[155] The main criterion in the definition of the graphematic word are empty spaces: "The graphematic word stands between two spaces and does not contain any spaces internally" (Fuhrhop 2008: 193, my translation).[156] This is a purely graphetic definition, but it is supplemented by three additional features that the '(proto)typical' graphematic word exhibits (cf. Fuhrhop 2008: 194): 1) it consists of one or more graphematic syllable(s) (as defined above in Section 2.4), 2) it is an unbroken sequence of graphemes, and 3) it contains a maximum of one majuscule word-initially.

155 „Der Begriff des graphematischen Wortes wird hier am Deutschen erarbeitet. Einige Aspekte sollten verallgemeinerbar sein, andere werden es nicht sein."

156 „Das graphematische Wort steht zwischen zwei Leerzeichen und enthält intern keine Leerzeichen."

The definition of the graphematic word as everything that stands between two empty spaces raises the question of whether punctuation marks – both sentence marks such as the period <.> and word marks such as the apostrophe <’>, which are not separated from letters by empty spaces but are enclitic to them, are treated as parts of graphematic words. For the word marks, i.e. the apostrophe <’>, the hyphen <->, the slash </>,[157] and the period after abbreviations <.>, the answer is yes, they are part of graphematic words (cf. Buchmann 2015; Evertz 2016: 392). By contrast, sentence marks such as the period used at the end of sentences are not analyzed as parts of the graphematic word. Fuhrhop (2008: 217) explains this by arguing that, for instance, <word.>, which is positioned at the end of the previous sentence, is merely a positional variant: if <word> were to appear sentence-internally, it could very well just be <word> or <word,>. As a sentence mark, the period <.> is enclitic to the graphematic words preceding it but is not a part of them, whereas the period after abbreviations as in <etc.> is an integral part of the word and occurs with it wherever the word appears within a sentence. A notable exception is the sentence-final position, where, depending on the analysis, the abbreviation period and the sentence period either merge or one of them is deleted (cf. Bredel 2008: 28). Evertz (2016: 391) provides an additional argument: treating <words.>, <words?>, and <words,> as distinct graphematic words would violate the principle of economy. Like Fuhrhop, he refers to Bredel’s (2008) theory of clitics and fillers to explain that sentence marks are not part of graphematic words, whereas word marks are (cf. Section 2.6).

The graphematic word is a concept that poignantly proves that writing is not dependent on speech or, more precisely, phonology. In this vein, Fuhrhop & Peters (2013: 251) stress that graphematic words differ markedly from phonological words, which are often much smaller units. Against this background, it is unsurprising that graphematic words frequently do not correspond with phonological words. By comparison, both the morphological and the syntactic words are crucial in the makeup of the graphematic word (cf. Fuhrhop 2008; Evertz 2016: 394). Even though these two types of words are most often congruent with graphematic words (and with each other), there are exceptions. Exceptions, too, as Fuhrhop (2008: 225) notes, are influenced by morphology and syntax: Consider, for instance, German verbs whose morphological subcomponents can be separated within the sentence such as <anfangen> ‘to begin’, where <an> can be separated from <fangen>, cf. <Er fängt morgen an.> ‘He starts tomorrow’. Syntactically, <fängt> and <an> are not separate words (admittedly, this also depends on the nature of the syntactic analysis) – graphematically, they are.

In the context of the graphematic syllable (cf. the previous section), it was argued that analyses based on visual criteria such as [±length] are not graphematic but graphetic. For the graphematic word, a purely graphetic analysis that is based on empty spaces would identify also the above-mentioned <word,>

157 In later works (e.g. Fuhrhop & Peters 2013), the slash is omitted in the description of word marks because it can be verbalized (cf. also Section 2.6).

as a graphematic word. Evidently, thus, an analysis that determines that <word,> should not be classified as a graphematic word transcends the boundaries of graphetics and is indeed a graphematic matter. Specifically, graphematic classifications become necessary such as a distinction between word marks and sentence marks. The rough basis of the analysis – empty spaces – is graphetic, but the necessary 'fine-tuning', so to speak, is indeed graphematic, illustrating the limits of a purely graphetic analysis.

When considering second-order empty spaces in the world's writing systems, i.e. those that visually demarcate one-dimensional graphetic sequences (cf. Section 1.2.2), it becomes evident that, at least synchronically, they are linguistically functionalized to separate morphosyntactic words in all alphabets and abjads, but only in some abudigas and in no morphographic writing systems. In systems in which these empty spaces do not indicate graphematic units that correspond with words, they – partially in combination with punctuation – make visible graphematic units that correspond with syntactic units. As for the graphematic syllable, for the graphematic word, too, we can ask whether there exist any salient visual indicators other than empty spaces that visualize graphematic units larger than the segmental graphemes. The next step would be to investigate whether these visually defined graphematic units correspond with the word (the morphological, syntactic, or even phonological word).

In the writing system of Japanese, the second-order empty space indicates mostly (but not exclusively) syntactic units. It occurs after non-Japanese punctuation marks which were borrowed from Western writing systems such as the question mark. In a purely visual analysis, a larger empty space that appears like a second-order empty space is also found after the period <。 >. According to most analyses, however, the impression of an extra empty space is created by the fact that the period occupies the bottom left corner of its own segmental space – notably, this means that unlike in Western writing systems, it is not enclitic to the basic shape that precedes it. Given the lack of empty spaces between 'words', a crucial cue for the segmentation of words is the alternation between the different component scripts. This alternation only works as a cue because it is not arbitrary but reflects the different graphematic functions the scripts fulfill: morphographic *kanji* are used for lexical morphemes (nouns, verbs, adjective stems, some adverbs) and syllabographic *hiragana* for particles, auxiliary verbs, inflectional affixes of nouns, verbs, adjectives, and so on,[158] to mention only the functions of the two central components (cf. Smith 1996: 209–212). Since most inflectional morphemes in Japanese operate at the ends of words, hiragana graphemes quite consistently mark this latter part of words. Consequently, a sequence of a hiragana grapheme followed by a kanji grapheme marks the beginning of a new word, with the former representing the end of the first word and the latter the beginning of the next. In

158 And syllabographic *katakana* is used for even more functions: "[...] to write foreign names and loanwords, onomatopoetic and mimetic words, exclamations, and some specialized scientific terminology" (Smith 1996: 212).

this vein, the writing system's users state that when only the syllabographic kana inventories are employed (which is sometimes the case in some special contexts such as literacy acquisition), this obscures word boundaries and consequently hinders the detection of words during the reading process (cf. Section 7.2.8). As one user – a native speaker and reader/writer of Japanese – put it in a related thread on the Reddit sub *Learn Japanese*: "If you're reading in all hiragana/katakana, it's very hard to understand the word boundaries. (In kids [sic] books, they tend to put spaces) Once you learn more kanji/grammar patterns, the word boundaries become much clearer".[159]

An interesting question is whether the visual cues provided by this script alternation in Japanese are in any way comparable to the [±length] feature that is prominent in Roman script and a number of other scripts (cf. the previous section). In those scripts, ascenders and descenders of basic shapes are visible even when the graphematic status of a basic shape has not yet been evaluated. In Japanese, by contrast, script alternation is not as much a visual cue as it is a graphematic one: while the kanji basic shapes look different than the kana basic shapes inventories, the visual differences between them are not straightforwardly and consistently attributable to the membership to their respective script. Indeed, it is questionable whether a kanji basic shape and a kana (e.g. hiragana) basic shape look more dissimilar than two given kanji basic shapes or two hiragana basic shapes do script-internally. Thus, whereas in writing systems using Roman script, readers can skim a line without actually reading it and still perceive and distinguish the long basic shapes from less salient non-long or 'compact' basic shapes, at least the ones at the extreme poles of the gradual length continuum, in Japanese, the alternation between scripts is not attributable to the fact that they appear visually different. Instead, it is simply a matter of knowing which script a basic shape belongs to. In short, is a matter of classification. I am not going into detail about whether this categorization is a graphetic or graphematic matter, as a person can potentially know that a basic shape x belongs to script y without knowing what linguistic unit basic shape x corresponds with exactly.[160] However, in any case, such a classification is likely a more conscious process than simply perceiving which basic shapes are long or compact in relation to each other and based on the vertical spatial organization of the line.

In the writing system of Chinese, in which only morphographic hanzi are used (comparable in function to Japanese kanji, which are of course derived from hanzi), words remain entirely unmarked. As Chinese has a morphosyllabographic writing system, every grapheme, with only very few exceptions, corresponds with a morpheme, and frequently, these morphemes are free and we could

159 https://www.reddit.com/r/LearnJapanese/comments/31u20m/getting_past_the_lack_of_spaces_between_words/cq4zd2v/ (June 2nd, 2020); comment posted by user *mirukushake*.

160 However, knowing which script a basic shape belongs to often entails knowing at least what *type of unit* it refers to, even if the exact graphematic value is unknown to a person.

classify them as ‘words’ or ‘lexemes’. However, given that synchronically, most words in Chinese are actually polysyllabic (cf. Yen et al. 2012: 1009), often, single graphemes constitute only one part of a polymorphemic word or compound. Therefore, in Chinese, the first-order empty space – the one between basic shapes (or, in a graphematic sense, graphemes) – does *not* indicate words in a transparent manner, and not even free morphemes, as bound morphemes are treated the same way. What this empty space does indicate is simply morphographic graphemes which indirectly, given the phonological representation of corresponding morphemes, also indicate phonological syllables. How users can still perceive and process words in Chinese and other systems lacking this type of empty space (e.g. Thai) will be discussed in Section 7.2.8.

## 2.6 Graphematic sentence and punctuation

The graphematic sentence is special. For its definition, no other definition of ‘sentence’ stemming from a different linguistic subsystem external to the writing system – such as syntax – serves as a basis. Thus, the relationship of dependence between visual units and linguistic units is seemingly reversed. Accordingly, as Schmidt (2016) argues, the graphematic sentence influences colloquial as well as linguistic understandings of ‘sentence’. This situation can at least partially be attributed to the lack of a consistent syntactic (or other) definition of ‘sentence’. At least in German, the sentence in a syntactic sense and the graphematic sentence do not necessarily overlap (cf. Schmidt 2016: 222). Take the graphematic sentence <Olivenöl, zwei Regale voll.> ‘Olive oil, two shelves full.’ which is ‘smaller’ than a syntactic sentence in that it lacks some of the features that a syntactic sentence must exhibit – most prominently a finite verb. Evidently, graphematic sentences can be both shorter – and lack some grammatical constituents – or longer than sentences in the syntactic sense.

Schmidt (2016: 219) points out that the existing definition of the graphematic sentence is inherently circular: claiming that it (1) begins with a sentence-initial majuscule and (2) ends with a sentence mark, which is true for most alphabets (although he exclusively refers to German), is problematic given that both of these criteria (just look at their designations) actually require a pre-existing definition of ‘sentence’. In his attempt at an inner-graphematic analysis, Schmidt (2016: 234) defines the graphematic sentence as a suprasegmental unit of writing, whereas, as explained in the context of the graphematic syllable, I prefer the term polysegmental (cf. Section 2.4). Schmidt aptly illustrates that readers need not know the linguistic values of graphemes that constitute a graphematic sentence to recognize it as such. Accordingly, in a structure such as <Xxx xxx xxxxxx. Xx xxxxx xxx Xxxx xxxx xx. Xxx, xxx Xxxxxx!>, three graphematic sentences can be discerned purely on visual grounds. Therefore, at this point in the analysis, in line with my earlier remarks, it is reasonable to question the adequacy of designating this unit *graphematic* sentence instead of opting for *graphetic* sentence. Note, however, that the empty spaces separating graphe(ma)tic sentences from one another do not differ visually from those that stand between graphe(ma)tic words, i.e. they are of the same size.[161] This means that the distinction must include information beyond the visual level. Thus, for the time being, it appears justified to adhere to the term graphematic sentence.

The two visual cues that are prominent in the structure provided by Schmidt are capitalization at the beginning of the structure and punctuation at its end. However, as argued above, for a purely graphematic definition of the graphematic sentence, these criteria do not suffice: they cannot determine a graphematic sentence since they are simultaneously determined by it. Crucially, even graphetically, they are insufficient to differentiate between the graphematic sentence and the graphematic word. For instance, in German, but also in other alphabets, sentence-internal capitalization is not uncommon. Furthermore, because of the polyfunctionality of the shape of the period <.> which can be used either as a word or a sentence mark, the middle sequence (in bold print) in <Xxx xxxx. **Xxx.** Xxxxx xxxx Xxxx xxx.> could either be an abbreviation (and, thus, a graphematic word) or a graphematic sentence.[162] Graphotactically, it appears, graphematic sentences cannot be defined in isolation but only in context. Further information required to recognize a graphematic sentence includes the end of the preceding structure and the beginning of the following structure. For instance, if a period precedes a capitalized word, it is highly likely that this instance of capitalization marks the beginning of a graphematic sentence. Also, if any word in this sentence is followed by a period which in turn is followed by both an empty space and a capitalized word, this possibly marks the end of the graphematic sentence. The corresponding

161 Note that in English, there exists a practice of placing two empty spaces after a sentence-final period, which of course equals an empty space that is larger than the one between words. Interestingly, the question of using one vs. two spaces after a period has users of English divided into two camps (‘one-spacers’ vs. ‘two-spacers’) and results in heated discourses – mainly on social media. Two key points of the debate came in 2018, when a psycholinguistic study suggested double spaces aid reading processes (cf. Johnson, Bui & Schmitt 2018), and in 2020, when the dominant text processing software Microsoft Word started flagging double spaces as mistakes (cf. https://www.theverge.com/2020/4/24/21234170/microsoft-word-two-spaces-period-error-correction-great-space-debate, June 4th, 2020).

162 Schmidt (2016: 242-246) discusses the problem of distinguishing the period after abbreviations, which is a word mark, from the one used sentence-finally, which is a sentence mark. He refers to borderline cases in which abbreviation periods behave exactly like sentence-final periods, i.e. occur in the same graphotactic context, cf. the period (and following majuscule) after <a> in <Robert Wiene u. a. Regisseure prägten den Expressionismus.> ‘Robert Wiene and other directors shaped expressionism’ (cf. Schmidt 2016: 243). This is especially problematic for abbreviations that are simultaneously well-formed graphematic syllables such as <sog.> (German abbreviation for *sogenannt*, ‘so-called’). Following Schmidt’s analysis, the period in <sog.>, if it is followed by a majuscule, would have to be interpreted as a sentence-final period. This leads to the concession that an ‘inner-graphematic’ analysis has its limits. Note that also in Bredel’s (2008, 2011) analysis of punctuation, a differentiation between the sentence period and the abbreviation period presents a problem.

graphotactic structure that Schmidt (2016: 234) arrives at for the transition from one graphematic sentence to the next one is [.][ ][X][x$^{\text{min/maj}}$].

This structure can be verbalized as follows: a period followed by an empty space followed by a majuscule followed by either a majuscule or a minuscule. In prototypical instances, the first majuscule is indeed followed a minuscule; notable exceptions are cases in which a sentence starts with an acronym such as <EU> or a word written in all caps. The sentence-final period is a placeholder for all sentence marks that visually exhibit a dot on the base line – this includes |!| and |?|, but also |:| (for a discussion of the colon in this position, cf. Schmidt 2016: 237–239). A more complex schema that allows for more marginal cases as well (e.g. parentheses at the end or the beginning of a sentence), reads: [.\$^{\text{fin}}$][][**$^{\textbf{ini}}$\X][x$^{\textbf{min/maj}}$]... [x][x][][x][x]...[.\$^{\textbf{fin}}$]**[][$^{\text{ini}}$\X][x$^{\text{min/maj}}$]. Here, the part printed in bold represents a single graphematic sentence (cf. Schmidt 2016: 248): \$^{\text{ini}}$ stands for an optional initial and \$^{\text{fin}}$ for an optional final cliticizing punctuation mark. The opening parenthesis <(> and opening quotation marks <“> occur sentence-initially while their closing counterparts <)> <”> are located sentence-finally.

Schmidt’s analysis emphasizes that both visually and functionally, punctuation is vital to the graphematic sentence’s definition. Indeed, it can be argued that punctuation not merely indicates sentences but *constitutes* them (cf. Schmidt 2016: 215, 247). This appears to be, if not an absolute universal, at least a universal tendency in the world’s writing systems. In every writing system, there is a second-order empty space, and in the ones in which it does not demarcate units that correspond with (morphosyntactic) words, it indicates sentences or other syntactic units, and it does so almost always in combination with punctuation. Of course, this is a broad generalization, and there are exceptions: Thai exhibits spacing between syntactic units – not only sentences but also clauses – and lacks periods that mark the end of a sentence (cf. Danvivathana 1987: 262, 269). However, second-order empty spaces are not only used between syntactic units, as additional orthographic conventions require it (such as ‘Put a space between a person’s military rank and their name’, cf. Wathabunditkul 2003); these more marginal instances of the empty space greatly complicate the overall picture.

Every alphabet in use today has empty spaces between both words and sentences. Additionally, every alphabet features some sort of punctuation. A small set of punctuation marks – including the period <.> and the comma <,> – are even more widespread than the most prominent and oft-adopted scripts (such as Roman, Arabic, Cyrillic) as they are found across a large range of writing systems regardless of the script they use. Conversely, the second criterion central to the graphematic sentence, capitalization, is not even exhibited by all alphabets. Take the Georgian alphabet, for example, which uses a unicase script. Korean Hangul, too, is a unicase script that is structurally alphabetical but, depending on the analysis, possibly functionally syllabographic (cf. Section 2.7). Arabic and Hebrew, two abjads, also lack capitalization. However, both use punctuation, and empty spaces occur between both words and sentences. Capitalization is alien to the writing systems of Thai, an abugida, the mixed syllabographic/morphographic Japanese system as well as the morphographic Chinese system. They all have punctuation, and empty spaces mark units larger than words.

With her seminal (German) works on the topic, Bredel (2008, 2011) changed the view of punctuation. Notably, concerning generalizability, she states explicitly that she focuses only on a comprehensive description of the German system and that a comparison of punctuation systems in different writing systems constitutes a separate task (cf. Bredel 2009: 118). The first step in that direction is the application of her ideas to the punctuation system of English (cf. Kirchhoff & Primus 2016). In line with what has been stated for the graphematic units described in the sections above, Bredel supports a methodologically autonomous approach to studying writing. Consequently, she switches the perspective in which punctuation is commonly analyzed. While most prior conceptions interpreted linguistic constructions (sentences, clauses) or features thereof (declarative, interrogative, prosodic features, etc.) as *inputs* or necessary contexts for punctuation, Bredel treats them as the opposite: as *outputs* of the reading process, a process in which the reader is guided by punctuation marks. The former, traditional view is what Bredel terms the *offline* view, while the latter view that she proposes is called *online* view (cf. Bredel 2011: 5). In short, the core of her analysis of punctuation is that punctuation marks serve as instructions for the readers that help them in navigating the reading process. This leads to the crucial implication that they occur only if readers need to deviate from the 'default strategies' of reading (cf. Bredel 2008: 18, 2009: 118).

Analytically, Bredel's approach differs from older, descriptive conceptions in that it puts language processing front and center. Notably, thus, her analysis is not structuralist but functional. Concerning the mentioned descriptive conceptions, one of the noteworthy connotations that German punctuation (and other similar punctuation systems as well) never quite lost is that it is (directly) associated with prosody, or more deterministically: that it depicts prosodic features. As a result, in the German tradition of describing description, so-called 'rhythmic-intonational' (cf. Baudusch 1976: 199) or 'rhetorical-intonational' principles (cf. Kirchhoff 2017: 19–22) were always of relevance. At one point, however, the analytical focus shifted away from intonation and onto syntax, the new consensus being that punctuation mainly functions to indicate syntactic units and relations (cf. Behrens 1989). These traditional approaches to punctuation underline the two different functions associated with punctuation: indicating prosody and indicating syntax. Notably, descriptions of German punctuation that outline a diachronic development away from the prosodic principle and towards the syntactic principle (such as the remarks above) imply that these principles are somehow mutually exclusive or that only one of them can be dominant while the other is of relevance only secondarily. Fittingly, in a description of several European punctuation systems (cf. Dokumente 1939), these systems are also classified as either prosodic or syntactic, underlining the view that there is a categorical division between the two. An additional aspect that enters the mix is stylistic freedom (cf. Nunberg, Briscoe & Huddleston 2002: 1727; Kirchhoff 2016: 399), which – as it concerns not only punctuation – I will mention in the discussion of *graphostylistics* (see Chapter 3). In a survey of the functions of punctuation, Kirchhoff (2016: 414) explains that the correlation between punctuation and prosodic features is not direct but rather a consequence of the connection between punctuation and syntactic features. In turn, it is the close ties between syntax and prosody that create the impression that punctuation was (directly) prosodic.

Without going too much into detail, several notable aspects of Bredel’s innovative approach need to be mentioned. First, as Kirchhoff & Primus (2016: 94) note, in comparison with other conceptions, Bredel (2011: 9) operates with a narrow set of twelve punctuation marks: < . ; , : - – … ’ ? ! ( ) „ “ >. This set results from a graphotactic analysis that allows a distinction between basic shapes of the categories letter, digit, special character, punctuation mark, and empty space (cf. Section 1.2.1). Bredel’s most remarkable achievement for a theory of punctuation is the discovery of (potentially generalizable) form-function correlations for punctuation marks, which resembles Primus’ (2004, 2006) analysis of the lowercase basic shapes of Roman script. For the formal description of the inventory of punctuation marks, Bredel posits three graphetic features: [±EMPTY], [±VERTICAL], and [±REDUPLICATED]. Punctuation marks that have contact with the base line, i.e. < . ? ! : ; ( ) „ “ >, are [–EMPTY] while < - – … ‘ > do not touch the baseline and are [+EMPTY]. Notably, Bredel works with historical forms of two punctuation marks to arrive at their synchronic feature values, namely the quotation marks < „ “ > that were formerly written as < ⟩ ⟨ > and are thus [–EMPTY] and the ellipsis mark < … > that formerly appeared as three strokes in the high subspace of the linear space (i.e. as < /// >) and is thus [+EMPTY] (cf. Bredel 2009: 120). These historical forms also explain how she arrives at the feature values for [±VERTICAL], where the deciding criterion is whether a mark occupies the high space or extends into it: thus, < ? ! /// ’ ⟩ ⟨ ( ) > are [+VERTICAL] and < - – : ; , . > are [–VERTICAL]. The final feature, [±REDUPLICATED], depends on whether the base element of a mark is visually reduplicated: < /// – ⟩ ⟨ ( ) : >[163] are [+REDUPLICATED], < - ; , . ? ! ’ > are [–REDUPLICATED].

Accordingly, every punctuation mark exhibits three feature values that allow the assumption of graphetic feature classes (cf. Bredel 2011: 4). The marks within these classes behave similarly with regards to function, underlining Bredel’s assumption of form-function correlations. Specifically, [±EMPTY] is of graphotactic relevance: marks that are [–EMPTY] such as the period <.> are characterized by the fact that their preceding and following ‘neighbors’ are never a basic shape of the same class: what precedes the period is a letter, digit, or special character, what follows it is an empty space. Marks that are [+EMPTY], on the other hand, can be preceded and followed by basic shapes of the same class, as in <geht’s>, where the apostrophe is both preceded and followed by letters.

Furthermore, as marks which are [–EMPTY] cliticize on and are positioned within the same segmental space as the unit that precedes them, Bredel

163 Bredel (2008: 29), again reasoning historically, interprets the dash <–> as a reduplicated hyphen: <-->.

(2009: 121) calls them *clitics*. Marks that are [+empty], on the other hand, occupy their own segmental space and are *fillers*. A functional difference between fillers and clitics echoes the distinction between graphetics and graphematics. Some visual units, such as graphematic words, but also paragraphs and columns, are constituted by the presence or absence of graphetic material as well as its organization on the writing surface – they are *graphically coded*. As such, readers *scan* them, i.e. perceive them immediately with their eyes. Crucially, words, paragraphs, etc. can be recognized as such even when readers have not extracted their content (cf. the graphematic sentence that consists of only x's above). By contrast, there are units that are not graphically coded: for example, graphetic material of different classes – the above-mentioned letters, digits, special characters – can be combined to larger units (such as <example> or <474>), and their content must be processed by the reader. These units are *linguistically coded* and are not perceived immediately during scanning – they must be *processed* (cf. Bredel 2011: 24). Based on this distinction, Bredel shows that fillers aid scanning, whereas clitics support processing or, in other words: fillers are relevant for lower levels of processing, i.e. physiology, and clitics for higher levels, i.e. cognition.

The remaining graphetic features are also functionally relevant. [±reduplicated] provides information about the scope of a punctuation mark: [−reduplicated] marks have the word and the sentence as possible scopes, while [+reduplicated] marks are located at the textual level. Finally, punctuation marks are (predominantly) relevant for one of two dimensions: the cognitive dimension (= parsing processes) or the communicative dimension (= which reader/writer roles are established through the marks?). Marks that are [−vertical] concern the former and [+vertical] the latter – for details, cf. (in German) Bredel (2008, 2011).

As mentioned above, many writing systems share at least a subset of the punctuation marks treated in Bredel's analysis. These shared marks quite possibly have largely similar functions in other systems, especially since 'instructing the reader' or 'guiding the reading process' are certainly not language-specific functions of punctuation. As for many grapholinguistic issues, more research – in this case writing system-specific analyses – is needed.

## 2.7 Writing system typology

This section provides an overview of the most relevant proposals for a typology of writing systems and collects pertinent open questions. Although abundant research has been carried out in this grapholinguistic subfield, leading to several well-known and accepted typologies, there is a lack of consensus as to which typology should be regarded as the most accurate one – partially also because there is no general agreement over what precisely a typology should achieve. Gnanadesikan (2017: 14) even goes as far as claiming "[t]he typology of writing systems is not a topic on which any two grammatologists appear ready to agree", with 'grammatologist' being her term for what I would call 'grapholinguist'. While the situation might not be as dramatic, it is true that the field is undoubtedly fraught with disagreements over several aspects. These divisive issues stem from the selective nature of typologies and typology-building that was already addressed in the context of a nascent typology of scripts in Section 1.3. In this context, Coulmas (1996a: 520) notes that "typologies are necessarily theoretically informed and selective[,] focusing on particular properties of writing systems rather than on others". The criteria that a typology is based on should be both informative and analytically valuable. By contrast, criteria such as genetic affiliation or geographic origin are not useful, which means terms like 'Chinese-derived writing' or 'Central American writing' are of no value typologically[164] (cf. Coulmas 1996b: 1381).

Most existing typologies of writing systems have focused on the smallest functional units of writing systems, most often defined as those units that represent or correspond with "the system's elementary signs, words or morphemes or syllables or phonemes" (Coulmas 1996b: 1381). Coulmas refers to these smallest written units as *basic units of operation*. This aligns perfectly with one of the three criteria in my conception of the grapheme: according to the linguistic value criterion, the grapheme corresponds with a linguistic unit. In this vein, for example, the first seven chapters in Daniels' (2018) comprehensive work on the world's writing systems are also based on linguistic units (syllables, segments, consonants, moras, clusters, morphemes, and words, in that order) and the writing systems that are based on those units. Despite the fact that this criterion clearly dominates writing system typology, it is by no means the only possible one, and it is in itself reductive, necessarily disregarding many other aspects. Some of these neglected aspects are listed by Coulmas (1996b: 1381):

> Typologies do not usually refer to higher-level organizational principles of writing, e.g., chapters, sections, paragraphs, and sentences by means of which text is segmentable, or properties of text such as direction (left, right), axis (horizontal, perpendicular) or lining (top to bottom, bottom to top). Accordingly, punctuation is generally disregarded in typologies of writing systems.

Gnanadesikan (2017: 14–15) likewise mentions some of the possible criteria besides the nature of writing systems' basic units of operation, and her list shows an overlap with Coulmas': "a set of signs, the spatial arrangement of the signs, […] and language-specific orthographic rules by which the signs are interpreted". As was already mentioned above, not all of these aspects can be accounted for in a typology, as it is necessary to prioritize some criteria at the expense of others (cf. Coulmas 1996b: 1386). However, this practice of devaluing most of the possible alternative (or additional) criteria in favor of singling out the role distinct linguistic units play as correspondences for graphemes in different writing systems has resulted

164 However, to completely disregard geography would be shortsighted, as it is possible that there exist areal phenomena that are also interesting to analyze from a typological perspective.

in an overall rough nature of typologies that simply overlooks many distinctions. This leads Weingarten (2011: 12) to his conclusion that

> [t]he typology of writing systems is still in its beginnings. The types proposed to date [...] may highlight certain basic characteristics of a writing system but they cannot, for example, elucidate the fundamental differences between the French and the Italian writing system, which both belong to the alphabetic type.

To arrive at a more fine-grained typology, Weingarten broadly suggests the comparison of writing systems, and crucially, not only of those writing systems that differ fundamentally – as is mainly done in the present approach – but, as his mention of French and Italian underlines, also very similar ones that are, in the typologies dominant today, classified as belonging to the same type, e.g. the alphabet (cf., in this context, comparisons of Scandinavian writing systems in Lindqvist 2001 and of Romance writing systems in Meisenburg 1996; cf. also Gronemeyer 2015 for an application of comparative graphematics).

Figure 20 illustrates some of the major typologies of writing systems that are also mentioned in Joyce & Borgwaldt's overview (2011) (for a more detailed description of these typologies and a number of useful figures, cf. also Voß 2003 – in German). The major distinction made in all of them is that between *phonographic* and *non-phonographic* writing systems.[165] For the latter category, different terms have been proposed, most prominently *logography* and *morphography*. Following Sampson's (2015) view that the basic units of writing may refer to morphemic units, but not polymorphemic units (such as words), I prefer morphography (cf. also Joyce 2011).

What is obvious from this collection of typologies is that overall, there are four major types of writing systems, and three of them are *phonographic*: syllabographic, segmental (phonemic), and a marginal, sometimes included subsegmental (featural) type. The fourth important type is simultaneously the second broad category mentioned above, *morphographic*. Daniels (1990, 1996, 2017, 2018) has proposed the most widely accepted subtypes of the phonographically segmental branch: in addition to alphabets, he assumes abjads and abugidas. These three segmental subtypes differ in the categories of phonemes that the written units represent: in an *abjad*, "the characters denote consonants (only). [...] In an *abugida*, each character denotes a consonant accompanied by a specific vowel, and the other vowels are denoted by a consistent modification of the consonant symbols" (Daniels 1996: 4, emphasis in original). The terminology is not self-explanatory but well justified. Analogously to *alphabet*, Daniels coined the terms by using the names of the first four units of respective representative systems of the two types:

165 The most important typology omitted here is Sproat's (2000: 142). He proposes a two-dimensional typology in which one axis specifies the type of phonography (consonantal, polyconsonantal, alphabetic, core syllabic, syllabic) and the other informs about the amount of logography, with writing systems such as English with no logography positioned at one end of the spectrum and systems such as Sumerian or Japanese in which logography is a constitutive feature located at the other end.

for *abjad*, the Arabic writing system (in order: a–b–ja–di) and for *abugida*, the Ethiopic writing system (a–bu–gi–da) (cf. Daniels 1990: 729f.).

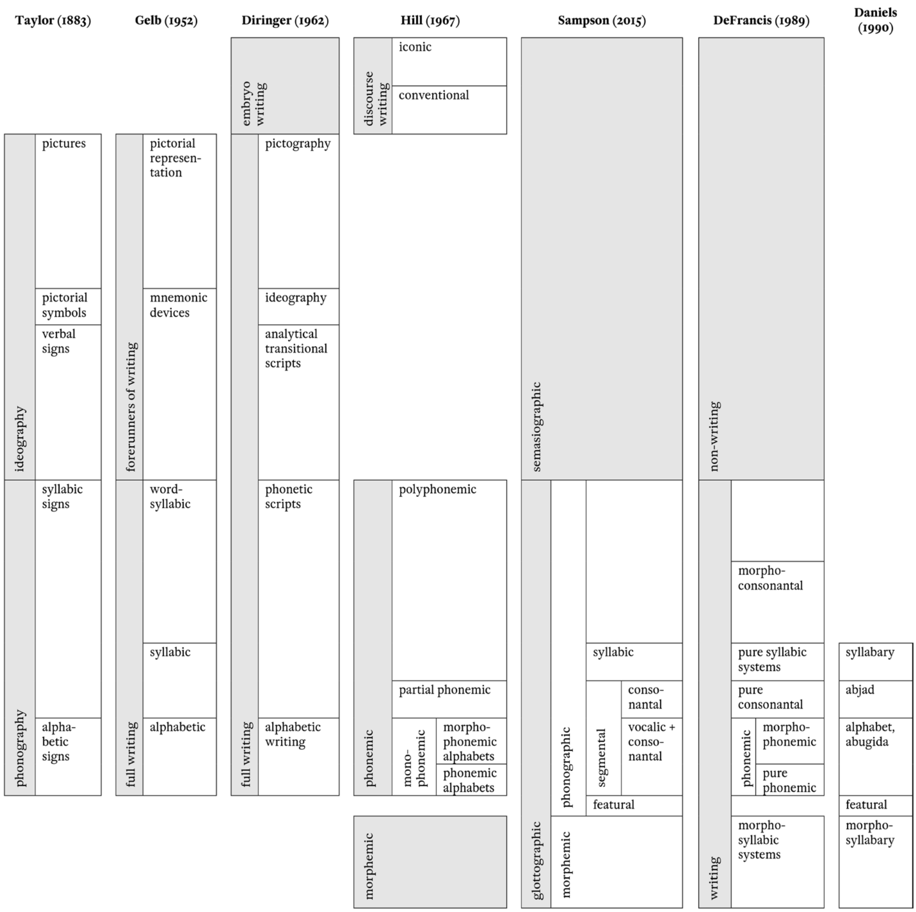



FIGURE 20. Prominent writing system typologies

Notably, Gnanadesikan (2017: 14, emphasis in original) echoes Weingarten's above-mentioned sentiment and argues that these types proposed by Daniels are not fine-grained enough, claiming that "a short list of simple one-word names, like *alphabet*, *abjad* and *abugida*, does not capture the full range of segmental scripts or the relationships between them". So, in her sound and impressive contribution, Gnanadesikan evaluates a number of categories (not unlike Faber's 1992) that lead to finer distinctions between types not included in the *alphabet–abjad–abugida* trichotomy. Table 4 lists her categories as well as the terms that she proposes for the types instantiated by the categories' different values. The resulting terminology is characterized by rather long designations for individual types. Although they, similarly to my proposed terms for the types of allography (cf. Section 2.3), appear cumbersome and do not make for elegant one-word designations such as Daniels', they are conceptually precise, which, from a scientific standpoint, is the favorable quality. Thus, in Gnanadesikan's terminology, an *alphabet* is a fully vowelled linear segmentary, an *abjad* is a consonantal linear segmentary, and an *abugida* is a mostly vowelled āksharik segmentary, although Gnanadesikan (2017: 32) notes that 'mostly vowelled', here, really means 'all vowels but one'. Even though I believe the more precise terms are valuable when speaking about individual systems, Daniels' terms *alphabet*, *abjad*, and *abugida* still work as partially vague superordinate terms that subsume a number of systems exhibiting individual differences and idiosyncrasies, although, as Gnanadesikan aptly illustrated, some systems lie outside of their scope.

Table 4. Typological categories and terminology, from Gnanadesikan (2017: 28), slightly modified[166]

| Category | Values | Term |
|---|---|---|
| Characters (basically) represent segments | Yes | Segmentary/Phonemic writing system/Segmental writing system |
| | No | Other (e.g., Syllabary) |
| Other structures represented (other than those in 'higher-order structures' below) | Features | Featural |
| | Moras | Moraic |
| | None | (omit) |
| Higher-order structure represented | Peak/margin | Āksharik |
| | Syllables | Syllabically arranged/spaced |
| | None | Linear |
| Inclusion of vowels | All | Fully vowelled |
| | Most | Mostly vowelled |
| | Some | Partially vowelled |
| | None | Consonantal |

For the writing systems of German and Arabic, for instance, Gnanadesikan's proposal offers unambiguous categorical assignments: German is a prototypical alphabet and thus a fully vowelled linear segmentary. Arabic is a partially vowelled linear segmentary. The Thai writing system is an interesting case. It is tempting to call it a mostly vowelled āksharik segmentary. However, only some but not all of the vowels in Thai are represented by graphemes that are secondary in that they are 'bound', i.e. cannot appear on their own, are smaller in size, and/or are located above/below consonant graphemes. The remaining vowel graphemes are indeed of the same size as the consonant graphemes and are positioned linearly on the base line, just like consonant graphemes. Thus, Thai represents a bit of a mixture

166 Since Gnanadesikan uses *script* in a different sense, I changed the term in the table to *writing system*. Conceptually, however, we mean the same phenomenon.

concerning the representation of vowels; it is a mostly vowelled āksharik/linear segmentary.

Since non-segmental phonographic as well as morphographic systems are not treated in Gnanadesikan's proposal, I want to discuss an extension of her approach that includes all types of writing systems.

Gnanadesikan's first category, segmental vs. non-segmental, is logically constitutive for all non-segmental types of writing systems. In this category, we find syllabographic writing systems as non-segmental systems which are phonographic and morphographic writing systems as non-segmental systems which are, obviously, non-phonographic. This, ultimately, results in yet another trichotomy: segmental vs. syllabic vs. morphemic. As was already established in the discussion of the graphematic syllable (cf. Section 2.4), syllabographic writing systems such as Cherokee or Vai are by default what Gnanadesikan calls *syllabically spaced*, since there are empty spaces between (graphetically segmental) graphemes and these graphemes correspond with phonological syllables. The same cannot be posited for morphographic systems. However, since frequently, Chinese is mentioned as the type's most prominent representative, and in Chinese, a single morpheme almost always has as its phonological representation a single phonological syllable, the overall type of writing system is inaccurately referred to as 'morphosyllabary'. It is true that the writing system of Chinese, with notable exceptions,[167] is syllabically spaced, but it is imperative to note that this is only an *indirect* consequence of the fact that it is actually morphemically spaced. The same, however, does not apply to the morphographic kanji of the Japanese writing system. Although they have originally been borrowed from Chinese, in the native *kun'yomi* readings, Japanese morphemes are mostly polysyllabic, making it uncommon that a morpheme corresponds to only a single syllable. Thus, only one part of the Japanese writing system – the syllabographic part, as materialized by the two kana scripts – is consistently syllabically spaced. The morphographic kanji part is not.

The other categories in Gnanadesikan's approach do not apply to either syllabographic or morphographic writing systems. For example, in her exclusive treatment of segmental writing systems, she also addresses only those 'higher-order structures' that can be represented within the graphetic segmental space but not beyond, i.e. not in the linear space that represents a concatenation of segmental spaces. Thus, the graphematic syllable as defined for German in Section 2.4, being a polysegmental 'unit', is not included. Such units – the graphematic word is another example – fall into the scope of what Coulmas (1996b: 1381) calls "higher-level organizational principles of writing" and mentions as a domain usually not addressed by typologies. However, as mentioned at the outset of this chap-

**167** Mair (2011) describes a number of Chinese characters that are monomorphemic but polysyllabic. One example is <圕> *túshūguǎn* 'library', where one character represents a trisyllabic morpheme. Mair discards the firm belief that Chinese is monosyllabic as a "myth of innate monosyllabism of Chinese language, and even of Chinese writing, a myth with which students are indoctrinated worldwide" (cf. also DeFrancis 1984; Behr 2018).

ter, the story of graphematics and writing in general is anything but done with the description of the segmental grapheme, which, as the written correspondence of the “basic unit of operation”, is merely the base criterion for existing typologies. Supported by the preceding discussions of the grapheme as well as larger graphematic units, I will modify existing typologies precisely by integrating this question of higher-level organization. The basis for my argumentation will be Table 5.

TABLE 5. A typology of writing systems inclusive of higher levels

| level | **writing systems** TYPES / *linguistic units* | | **Korean** (FEATURAL) ALPHABET | **German** ALPHABET | **Thai** ABUGIDA | **Arabic** ABJAD | **Japanese** (kana) SYLLABARY | **Chinese** MORPHOGRAPHIC |
|---|---|---|---|---|---|---|---|---|
| phonological | *feature* | | [x] | [x][5] | | | | |
| | *phoneme* | *C* | [X] | X | X | X | | |
| | | *V* | [X] | X | x[1] | (x), X:[2] | | |
| | *syllable* | | X | [X] | [X] | [X] | X | X[3] |
| morphological | *morpheme* | | [X] | [X] | [X] | [X] | [X] | |
| | *word* | | X | X | [X] | X | [X] | [X][4] |
| syntactic | *sentence* | | X | X | X | X | X | X |
| textual | *larger units* | | X | X | X | X | X | X |

## Legend

X = a linguistic unit is represented in writing

x = a linguistic unit is represented in writing by a graphematic unit that is *graphetically* and/or *graphematically secondary*

( ) = the graphematic representation of this linguistic unit is *optional*

[ ] = a unit that is *not* made visible by empty spaces, i.e. does not fulfill the *empty space criterion*

= linguistic units that are the basic units of representation for *graphemes*

= these linguistic units are *not* represented in a writing system

## Notes

**1** There is a lowercase x in this cell because in Thai, like in other abugidas, the vowel graphemes are dependent on consonant graphemes, meaning they cannot occur alone. In some cases, they are also graphetically secondary as they are relatively smaller in size and attach to the consonants inside their segmental spaces, i.e. they are bound. In Thai, however, there are also vowel graphemes that are equal in relative size to consonant graphemes and occupy their own segmental space (see above).

**2** In Arabic, long vowels are always graphematically represented, which is signified by an uppercase X followed by a colon, the latter of which indicates vowel length in the IPA. Short vowels are commonly not represented, although they can optionally be written in some contexts such as material for L1 reading acquisition or teaching material for Arabic as L2. In case they are written, they are graphetically secondary to consonant graphemes, both in their placement above or below consonant graphemes as well as in their size, which is why the x in parentheses is lowercase.

[3] In most cases, morphemes in Chinese have monosyllabic phonological representations. Thus, although syllables are graphematically represented in Chinese due to spaces between graphemes, the correspondence between graphemes and syllables is only indirect as it is a consequence of the correspondence with morphemes. This cell in the table is occupied only because we are adopting the analytical direction of *writing → language*. Given the vast number of homophonous syllables in Chinese the opposite direction would not be feasible: while morphemes have a uniform correspondence with graphemes, a given phonological syllable (crucially without an associated meaning) can often correspond with numerous graphemes.

[4] In some cases, in Chinese, a single morpheme, if it is free, can already represent a word. If this is the case, the two units of morpheme and word merge visually and, consequently, words are graphetically represented through preceding and following empty spaces. However, many words in modern Chinese are polymorphemic and, thus, graphematically represented by more than one grapheme. These polymorphemic and, in turn, polygraphemic words are not separated from other words by specific preceding and following empty spaces and, thus, the word in its 'common sense' is not a unit that is graphetically or graphematically marked in Chinese.

[5] Primus (2004, 2006) postulates a form-function correlation for subsegmental components of lowercase basic shapes in Roman script, i.e. that certain features of elementary forms (such as visual length) correspond with phonological features such as place of articulation. This will be discussed in Section 6.2.

In Table 5, for each type of writing system, one representative system is included. In the table, they are listed with Daniels' terms, partially for reasons of formatting reasons, i.e. limited space inside table cells. This typology, now, illustrates not only that the "basic unit of operation" is the unit of correspondence for the writing systems' respective *graphemes* but also indicates those higher-level linguistic units that have a graphematic representation. This confirms that the *strict layer hypothesis*, transferred by Evertz (2016: 392) from phonology to graphematics, actually holds not only for the alphabetic German writing system but for all types of writing systems: accordingly, in a suprasegmental hierarchy, here understood as a polysegmental hierarchy (for the difference, take a look back at Section 2.4), a given unit must always consist of a combination of units of the immediate lower level. Simultaneously, it serves as a constitutive part of higher-level units. Thus, the graphematic representation of phonological syllables, if it is achieved polysegmentally, i.e. with more than one grapheme, must, at a lower level consist of graphemes, while at a higher level it forms part of the graphematic representation of morphemes. Another aspect that is emphasized in this typology is the interaction between graphetics and graphematics, specifically the central question of which graphematic units are simultaneously graphetic units – in other words: which graphematic units are visually salient.

The cells that have a dark grey background in Table 5 reveal which linguistic units a given writing system's graphemes correspond with. In the segmental writing systems, fittingly termed *segmentaries* by Gnanadesikan, these are phonemes, and they exhibit precisely the fine-grained distinctions that Gnanadesikan underlined: German and Korean are fully vowelled, Thai is mostly vowelled, and in Arabic, only one type of vowels is commonly graphematically represented. Notes (1) and (2) above give more information on the special situation in Thai and Arabic – in short, the minuscule x's, in contrast to the majuscule Xs, showcase that vowel graphemes are secondary while consonant graphemes are primary. In the kana parts of Japanese, graphemes correspond with syllables, and in Chinese, graphemes correspond with morphemes and, due to the abovementioned morphosyllabic nature of Chinese, simultaneously syllables, cf. note (3). This is not the case for Japanese kanji, which are not included in the table; as graphemes, they – at least in native readings – prototypically correspond only with morphemes but not necessarily single syllables.

Table 5 reveals yet another almost universal trait of writing systems: graphemes, as the smallest linguistically functional units, are simultaneously the smallest freestanding visual units that occupy segmental spaces and are constituted by the smallest empty space, the empty space between basic shapes. This is marked in the table by those Xs (or x's) that are not enclosed by square brackets and are located at the lowest level. The first notable exception, here, is Arabic, in which the basic shapes occupy segmental spaces but are mostly connected. Therefore, there is no consistent empty space between all basic shapes throughout the script. The smallest empty space does become visible, however, between the six basic shapes that are not connected to the left: <ا د ذ ر ز و>. If this empty space is visible, it is relatively smaller than the empty space between words. The second notable exception here is Korean, in which there are graphemes for the correspondence with both vowel and consonant phonemes, technically rendering Korean an alphabet. However, as Gnanadesikan also notes, these graphemes, which are themselves subsegmental, are syllabically arranged. Thus, it is not the consonant and vowel graphemes but the syllable blocks they compose in combination that occupy segmental spaces. Korean, thus, is the only example of a writing system in which the segmental space is consistently not filled by graphemes but a unit of writing that corresponds with a phonological syllable. However, as the *strict layer hypothesis* still correctly predicts, these written syllables consist of visually subsegmental vowel and consonant graphemes, which can be extracted transparently from the syllables, supporting the claim that Korean is not to be regarded a syllabary, at least not structurally. What it constitutes functionally is a different story (cf. Coulmas 2016: 45).

It is also noteworthy that in Korean, there exists another dimension to the subsegmental graphemes, the so-called featural dimension (cf. Sampson 2015: 143–166; cf. also Daniels 1990). 'Featural' refers to the fact that the basic shapes that materialize the subsegmental graphemes iconically depict the place of articulation of the phonemes they correspond with (cf. Figure 21). This visual iconicity (arguably a form of pictography) is diagrammatically shared (cf. Section 6.2) by basic shapes that have similar graphematic functions (such as |ㄱ|, |ㅋ|, and |ㄲ|, which are all used for graphemes that correspond with velar phonemes). However, this iconicity can neither be 'extracted' as a unit since it is rather an integral visual feature of the basic shapes nor does it consistently correspond with a linguistic 'unit' but instead only with a linguistic feature, specifically a phonological feature. As such, this featural dimension could be compared to the subsegmental components in Chinese that contribute a graphematic function to the overall grapheme but are not independent graphemes themselves.

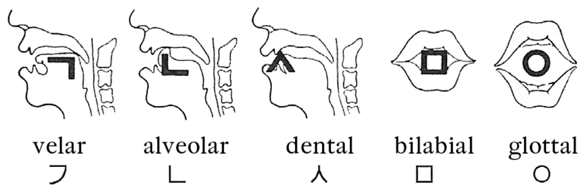



FIGURE 21: The featural level in Hangul, adapted from http://4.bp.blogspot.com/-rgIZ4WHh-v2o/VgNQhffUypI/AAAAAAAAMlI/5yLODu3qxlw/s400/consonant_1.jpg (May 28th, 2020)

Interestingly, above the level of graphemes, which corresponds with different linguistic levels in different types of writing systems, the next level that consistently produces visually salient graphematic units is hierarchically a rather high level: syntax. Thus, syntactic units such as clauses, phrases, and sentences (however they might be defined in a given context) are consistently marked by empty spaces and/or punctuation in the world's writing systems, as are various "units" such as the paragraph at an even higher linguistic level, the textual level.

In a nutshell, the rough working typology presented here captures many of the findings of the chapters on graphetics and graphematics and will, in the following, serve as a backdrop for the explanation-based part of this book in which various (types of) writing systems will be embedded and evaluated within a functionalist theoretical framework.

## 2.8 Epilogue: Dependency and autonomy revisited

> An autonomous graphematic analysis – in fact, an oxymoron – can only be undertaken tentatively and provisionally. In this case, one acts as if graphematic words, syllables, or letters can be decomposed into their material features and analyzed in their own right without having recognized them as meaningful units of a language and subjected them to this general point of view. If these relationships are kept in mind, however, this at least avoids the risk of positivist reductionism. Graphematic analysis is constitutively possible only as a *relatively* autonomous analysis. (Schmidt 2018: 47, my translation)[168]

168 „Eine autonome graphematische Analyse – eigentlich ein Oxymoron – kann nur versuchsweise und vorläufig vorgenommen werden. Man tut dann so, als könnte man graphematische Wörter, Silben oder Buchstaben in ihre materiellen Eigenschaften zerlegen und an und für sich betrachten, ohne sie je schon als sinnvolle Einheiten einer Sprache erkannt und damit diesem allgemeinen Gesichtspunkt unterworfen zu

After the graphematic module of writing systems – its structure, systematics, and units – was described and discussed extensively in the preceding subsections, let us return to the question raised in the prologue and arrive at the following (still preliminary) answer: graphematics is dependent on phonology, at least sometimes.

As Schmidt states in the quote above, "autonomous graphematic analysis" is actually an oxymoron, as graphematics by definition treats the linguistic functions of writing systems, i.e. the links between the visual and the linguistic. He accurately posits that in a graphematic analysis, it is impossible to evaluate units such as the "graphematic syllable" or the "graphematic word" without previously having identified them as units of a given language – as is also highlighted by their designations. These units are units of a given language precisely *because* their visual substance is linked to levels such as the syllabic or the morphological levels of a language. /haʊs/ and <house> are not two different words of different language systems, but units of different modalities. <house>, as a written word, does not refer to an extralinguistic referent (a specific existing house, for example) but to a linguistic structure. While it can be analyzed independently, this written word does not represent linguistic structure independently of the phonological representation /haʊs/, as is implied by regular grapheme-phoneme correspondences. Similarly, <木> *mù* 'tree' is connected to the signatum of the Chinese (specifically Mandarin) morpheme for {tree} which has /mù/ as its phonological representation. Chinese graphemes, and in general, the Chinese writing system, are not independent of the morphological system of Chinese.[169] It is a graphic linguistic system (which is actually just a paraphrase for "writing system") that represents the Chinese language. By contrast, the English writing system and its units correspond with the phonological level, but not exclusively. Other principles are at work as well, and morphological, lexical, etymological, etc. information can be and is represented at the expense of phonological transparency. However opaque an alphabet becomes, in its essence, it is still phonographic, which is where the ominous attribute *relative* comes into play.

In the prologue to this chapter, it was argued that a lot of what makes writing unique in comparison to speech is due to material properties, which are studied by graphetics. Graphetics, by way of its definition, is almost completely independent of language.[170] It subsumes much of what has been termed "in-

haben. Bleiben aber diese Zusammenhänge im Blick, ist zumindest die Gefahr eines positivistischen Reduktionismus gebannt. Die graphematische Analyse ist konstitutiv nur als *relativ* autonome Analyse möglich" (emphasis in original).

**169** However, they are independent of the phonological level, which is why the claim "graphematics is dependent on phonology" is accurate only for some writing systems, as in this case, graphematics is rather dependent on morphology. In short, in every writing system, there is a primary correspondence with a specific linguistic level.

**170** Graphetics is defined as the study of the materiality of writing, and the categorical distinction between writing and drawing, for instance, can only be made when writing is defined as the representation of language. Thus, when the narrow definition of writing as glottography is adhered to, graphetics by definition studies the materiality of a linguistic phenomenon. This, by extension, also makes it a linguistic discipline.

ner-graphematic analysis" in the German grapholinguistic tradition. By contrast, graphematics alone, especially if it is – from a methodological point of view – assumed to be a subsystem of language analogously structured to phonology, lacks the necessary tools to adequately study the materiality of writing (cf. Wehde 2000: 48). As the discussion about the graphematic syllable underlined, the assumption of "graphematic" units based on material features such as ascenders and descenders leads to categories specific to given scripts or writing systems, which are of little value for a more universal grapholinguistic theory and generally comparisons of writing systems. It also puts too much weight on the seemingly autonomous graphematic features of such a unit: it is true that there are systematic and significant points at which the graphematic syllable in German deviates from the phonological syllable, but the regular correspondences between these two units are still predominant. The same applies to the controversial definition of the grapheme: the way it has been autonomously defined in German grapholinguistics, it is a minimal lexical contrast in writing that is parallel to minimal lexical contrasts in phonology. This definition is restricted to segmental phonographic writing systems in which the function of being (minimally) lexically distinctive can be interpreted as parallel in writing and speech. Accordingly, the meaningfulness of this definition suffers severely from an attempt to extend it to other writing systems – which, in itself, does not sound like an unreasonable endeavor. In short, graphemes are not minimal lexical contrasts in writing that just happen to correspond to phonemes. This correspondence with phonemes (or syllables, morphemes, …) is constitutive of graphemes. Sometimes, for a variety of reasons, be they historical, political, etc., this correspondence becomes more opaque, giving rise to the impression that a writing system's graphematics is autonomous. Indeed, if the graphematic solution spaces of words in a writing system are large, it *will* develop its own strategies of dealing with ambiguities. Several systematic restraints, thus, will truly be inherent to the writing system, i.e. they will *not* be explainable with recourse to other linguistic subsystems but instead only in the context of an "inner-graphematic" analysis.[171] Orthographically determined spellings, which are regulated externally, are likewise often not explainable through recourse to other linguistic levels (and sometimes not even inner-graphematically, cf. Chapter 3). Writing, thus, *is* undeniably relatively autonomous, whether through its graphetics, its inner-graphematic systematics – including graphotactic restraints and types of allography –, or its orthographic regulation. However, in the view argued for here, this autonomy *follows* dependence. In the phonology-independent view, writing is initially treated as an autonomous system, and only *then* are correspon-

And while graphetics does not study the denotative functions of writing, it investigates the connotative functions of its materiality (cf. Chapter 1), which are, in a broad sense, also linguistic functions. This means graphetics is indeed only 'almost' independent of language.

**171** Cf. also Handel (2013: 24), who claims for the Chinese writing system: "[…] some internal elements can only be understood in purely graphic terms as units of a structured system that is independent of the spoken language".

dences between writing and other linguistic subsystems established and studied. This does not correspond with the phylogenetic relationship between writing and language: writing is *dependent* on language since writing represents language, but in the next step, it has developed autonomous features that imperatively need to be studied independently. These autonomous features, however, are much more pronounced in segmental phonographic writing systems. In other words, autonomy arises predominantly in the course of analyzing segmental writing systems and only because it is highly likely that segmentality was, in fact, provided by writing, i.e. that the segmentation of the continuous acoustic stream of speech into phon(em)es is an epiphenomenon afforded by writing. In this view, it was always the phoneme that was dependent on the grapheme and not the other way around (cf. Faber 1992; Davidson 2019). Non-segmental writing systems, by comparison, are based on very stable linguistic units such as the syllable or the morpheme, which are not constituted (or made more salient so that humans become aware of them) by writing. This may be the reason why there aren't as many "autonomous" graphematic structures in these systems (or they have just not been described yet).

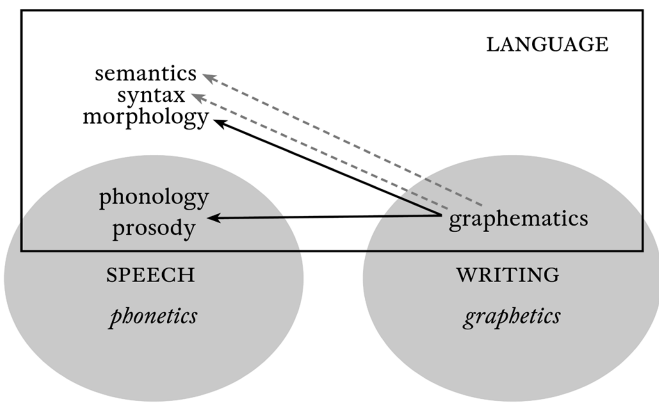



Figure 22. Conception of the relationship between speech and writing

# 3 Orthography

Let us turn to the last module of writing systems, the orthographic module. First, it is important to note that similar to *grapheme*, the term *orthography* is fraught with previous misuses, although in this case, the most prominent misuse is not a matter of controversial debate but instead the accepted reading: in English, *orthography* is most often used with a descriptive meaning, as a synonym of *writing system*.[172] A search for English literature on orthography, thus, will yield predominantly works dealing with writing systems descriptively, neglecting standardization of writing as a domain of language policy. This problem is also ostentatiously reflected in the practice of labeling the creation of writing systems for unwritten languages 'orthography development' (cf. Lüpke 2011) instead of the more accurate 'writing system development'.[173] In German (grapho)linguistics, *orthography* and *writing system* designate different phenomena. Orthography – as aptly illustrated by its etymology, consider Greek *ὀρθός orthós* 'right, true (*also*: straight, erect)' – is the standardization of a writing system (cf. Kohrt 1990: 116). It deals with the question of how to write (or, to use an inherently prescriptive term, spell) correctly with respect to external and explicit norms. It is not to be conflated with the internal and implicit regularities that reveal themselves in the use of a writing system and its resources, which are studied by graphematics (cf. Dürscheid 2016: 128) or, if they concern exclusively the material resources of writing, graphetics.[174] As Neef

---

**172** Neef (2005: 8) remarks that in the English-language literature, the distinction between *orthography* and *writing system* is largely unknown. Note that this could be due to the fact that the self-organizing English orthography (cf. Berg & Aronoff 2017) itself is not regulated in the same way as German orthography, for example, as there is no official codification and no official external regulator making orthographic decisions.

**173** In this context, however, one must admit that in the creation of new writing systems, a process that commonly involves linguists, writing systems are often consciously devised in a way that automatically minimizes the size of the graphematic solution space and, thus, variation. Therefore, they might not require a great deal of standardization. As these systems are artificially created for immediate use, the actors involved in the creation might not only have in mind *how to write* but indeed *how to write correctly*. Thus, from the outset, it might actually be an orthography that is created rather than an unstandardized writing system without an orthography.

**174** Admittedly, there is an additional, looser interpretation of *orthography* that stands between graphematics and the reading of orthography that is adhered to here. It concerns conventions that have developed out of the continuous use of a writing system in a literate community as they have been negotiated by users (and exist among users). These conventions need not be (externally and explicitly) codified, but that does

(2015: 715) points out, orthography is not an obligatory but an optional module of writing systems. Although nowadays, most systems are equipped with this module, in theory, writing systems can do without it (such as the written form(s) of Swiss German) – and have done so in the past. Ontologically, the orthographic module developed later than the graphematic module, as is underlined by Sebba (2007: 33), who notes that "the idea of a 'wrong' spelling is only two centuries or so old". Ironically, however, in literate communities in which the writing system in use exhibits an orthography, it is, in contrast to the graphematic module, phenomenologically primary. As Schmidt (2018: 34f.) argues, the graphematic module is not the underlying basis of an orthography – rather, the graphematic module is located *within* the orthography, so much so that even linguists frequently deal with graphematics only through orthography. In a literate community such as the German-speaking community in Germany, children do not learn to write – instead, from the beginning, they learn to write *correctly*. Writing, as a cultural technique, is almost always intricately and inseparably linked to normativity.

The reason that *orthography* came to be used as a general descriptive term is that it is not straightforwardly clear where to draw the line between graphematics and orthography. This makes the terminological misuse partially understandable. However, it remains untenable. As described above, graphematics deals with the relations, correspondences, and regularities between visual (or tactile) units and linguistic units. A writing system offers its users certain resources and possibilities, all of which they can, in principle, use to compose written utterances. The fact that there is seldom a one-to-one correspondence of basic shapes and linguistic units (or, more generally, graphic resources and linguistic functions) in writing systems is reflected by the existence of the above-mentioned graphematic solution space. It is the variation therein that calls for standardization.

The main (though not sole) function of graphematics is *communication*: it allows writers to compose messages that can be read – deciphered and understood – by potential addressees (which, in cases like shopping lists, include themselves). However, a given written unit, e.g. a written word, cannot be written randomly, as the graphematic solution space has its limits. In other words, a given possibility of spelling a word cannot violate too many graphematic relations – which are either phonographic or morphographic – in order to remain understandable: in the English alphabet, the word 'write' could, in theory, be spelled 'ryte' or 'right'[175] but not 'groeqx' (but cf. Sampson 2018). An orthography obliges the writer to obey the prescriptive rules that a community of writers has – more or less bindingly (see below) – agreed on by singling out one (or multiple) possible

not mean they do not exist, i.e. writing systems that do not exhibit an external and explicit orthography might still be 'orthographically' regulated, with rules being implicit and internal (cf. also Section 3.2).

175 This specific spelling would lead to the same phonological representation, but a different meaning as it is already used for an existing word with a different meaning.

spelling(s) as correct (cf. Karg 2015: 5). Accordingly, the functions of an orthography are *conventionalization* and *standardization*.

While this might clarify what an orthography is, it leaves open the question of how orthography diverges from graphematics in systems that feature an orthographic module. To answer this, the relations between the system, the norm, and the use need to be examined. As illustrated in Figure 23, these three realms overlap significantly but also display distinct areas.

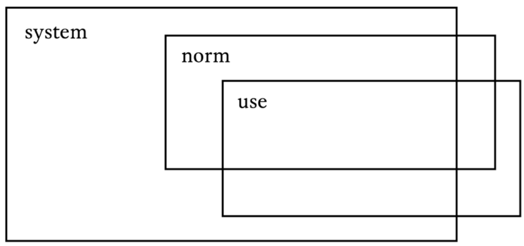



FIGURE 23. Triad of system, norm, and use, from Mesch & Noack (2016: 4)

First, a system must develop. This has occurred independently in at least three cases in which literacy was conceived from scratch: in the creations and the ensuing development of the Chinese, Mayan, and Sumerian writing systems. Commonly, however, writing systems are devised through the adoption of an existing (type of) writing system.[176] Usually, once established, writing systems are consistently in use. This use both constitutes the system and draws on it – it represents the "sum of the relatively synchronized internal norms of the writers of a language" (Schmidt 2018: 36, my translation).[177] If there is no system, there can logically be no use of it. And without use, arguably, a system cannot arise in the first place; and once it has been constituted, when it ceases to be used, a system meant for communication will die (cf. Section 8.2). Notably, the use of a system is dynamic in that users regularly introduce deviations in the form of spellings that are unsystematic, i.e. do not follow from the current graphematic regularities of the system. Here, a careful distinction must be made between conscious deviations – e.g. for reasons of style, creativity, innovation, or simply to reject the current norm[178] – and un-

176 Today, this appears to be the sole way of devising new writing systems, as the 'natural' independent creations of writing systems – unsophisticated ancient grammatogenies – worked only because the cultures in which they occurred had been primary oral cultures (cf. Ong 2012), i.e. did not know about writing (which, at least prior to the historically first invention of writing, did of course not exist).

177 „[…] Menge der relativ miteinander synchronisierten internen Normen der Schreiber einer Sprache […]."

178 In the German-language area, there are still people who do not conform to the reformed orthography of 1996/2006. Among them are also university professors, some of whose research interests even include orthography (examples are Utz Maas, Theodor Ickler). In this case, the deviance from the new standard is not to be regarded as an error or a mistake, but as an intentional choice – and, as such, a form of sociolinguistic action.

conscious deviations in the form of mistakes or errors. Note that deviations from the system are much rarer than deviations from the norm (see below).

Although they are (initially) perceived as deviations, unsystematic spellings can – depending on factors such as frequency of use – instigate a change of the system and eventually become part of it. However, changes of the system do not have to originate in use, as another process is central: standardization, which leads to the third part of the triad: norm. Ideally, only spellings that are part of the system should be part of the norm. For the instances for which this is indeed the case, Neef (2015) speaks of *systematic orthography*. However, spellings which are considered as orthographically correct sometimes do not conform to the regularities of the graphematic module. In other words, they are excluded from the graphematic solution space. They might still be used by users and gradually integrated into the system, but in extreme cases, spellings might neither be part of the system nor present in actual use. However, just as use can change the system, norms can change it too: an initially unsystematic spelling that is nonetheless codified as orthographically correct might come into (increasing) use as users are susceptible to normativity and prescriptivism. Ultimately, this can also lead to a change of the system. In sum, the diachronic development of writing systems can only be understood and explained through the triad of system, use, and norm (cf. Mesch & Noack 2016: 4).

This also helps to separate graphematics from orthography: the scope of graphematics is the system and its use, orthography focuses on the norm. As implied above, whereas orthographic rules are explicit and externally codified, graphematic 'rules', which are constituted by regularities, are implicit and internal and manifest themselves only in the use of the system (cf. Fuhrhop & Peters 2013: 186). However, these two types of rules are intricately related to one another: internal norms (which should probably be more appropriately called conventions) that are derived from and simultaneously show themselves in the use of the system can be externalized and become codified norms, and vice versa, external norms can be internalized (cf. Berg 2016a: 19).[179] As Schmidt (2018: 29) notes, the orthographic norm that is taught in the course of institutionalized literacy instruction in schools shapes the internal norms of writers, which, consequently, affects the use and with it the system, which Schmidt defines as the sum of empirical regularities. That way, orthography is effectively primary and greatly influences the graphematics of a writing system. This poses a problem in the analysis of written texts as products of the use of the system: it is most often impossible to tell internal rules and external rules apart as the text does not reveal which of them the writer followed (cf. Kohrt 1987: 341). However, Berg (2016a: 20) describes an interesting exception: if there is a deviation from the external, i.e. orthographic norm, but this deviation occurs consistently and displays its own systematics, then an underlying internal rule must be responsible. In this case, a deviation from the external norm

179 Kohrt (1990: 118) goes further and claims that it is not only possible that external norms are internalized but that an internalization is actually demanded.

is not the result of mistakes in performance, but due to a different system underlying competence. This means that a writer's underlying writing system may exhibit idiosyncratic (and, in a way, idiolectal) regularities that conform neither to the actual, inter-individual graphematic module nor its standardization, its orthography.

Notably, the graphetic module, too, is subject to orthographic regulation. As Kolers (1983: 383f.) remarks, "[it] is commonplace to worry about the rules of orthography as they affect reading, and many nations alter the spelling of words to bring spelling into line with current pronunciation" and asks whether, similarly, "the actual formation of characters can be put to principled test". Strikingly, orthographic regulation is much more uncommon in the graphetic module. This is probably because graphetics is visible, it is tangible, unlike graphematic relations which are somewhat "invisible" since they are abstract and materialized only by graphetics. Alterations of the graphetic module (such as Chinese character simplification) might, for example, be interpreted as more invasive since they change the visual appearance of writing. Also, like common, graphematically oriented orthographic reforms require updates of dictionaries, textbooks, etc., orthographic reforms that interfere with graphetics would require an adjustment of all technology suited for a given script. Yet, there do exist ways in which graphetics is orthographically regulated (cf. Sections 5.2, 7.3, and 8.3).

In the following, orthographies will be first investigated from a structural point of view before the focus is shifted onto the crucial sociolinguistic implications they have. To escape Eurocentrism, I will attempt to integrate orthographies of non-alphabetic writing systems into the discussion, although they have been widely neglected by the literature. In any case, it is important to note that most of what will be described here applies to orthography in general, regardless of the writing system it regulates.

## 3.1 Features of orthography

Nerius (2007, 2020) lists four features of orthography: as already mentioned, an orthography is an (1) external, codified norm. It is characterized by its (2) bindingness, its commonly (3) small degree of variability, and, even though it is prototypically static in nature, its (4) possible changeability. In the following, each of these features will be discussed in detail.

An orthography is an (1) *externally codified standardization*. Orthographic rules are codified in a set of regulations, in a dictionary, or – frequently – in both. As will be shown below, orthography is a central issue of language policy. As such, it is not seldom regulated by certain institutions with linguistic authority. In the pluricentric German-language region (encompassing Germany, Austria, Switzerland, but also Liechtenstein and parts of other countries), this is the *Rat* *für deutsche Rechtschreibung* (Council for German Orthography);[180] in Japan, it is the *bunkachō* (文化庁, Agency for Cultural Affairs),[181] a part of the Ministry of Education, Culture, Sports, Science and Technology; in Thailand, it is the *Ratchabandittayasapha* (ราชบัณฑิตยสภา, Royal Society of Thailand);[182] in South Korea, it is the *Gungnip Gugeowon* (국립국어원, National Institute of Korean Language),[183] to name a few examples. Sometimes – as is the case in Thailand – the set of orthographic regulations as well as a prescriptive dictionary (in this case the *Royal Institute Dictionary*) are published by the same regulator, which means the double codification of orthography is attended to by the same agent. Even when these two tasks are carried out by different stakeholders, there is a tendency for them to overlap: The *Duden*, for example, who publishes the most influential German dictionary, offers recommendations in that it designates 'preferred spellings' in cases in which two or more spellings for a word are deemed correct by the Council for German Orthography. In the case of <ph> vs. <f> in words such as <Typographie/Typografie> 'typography', the Duden explicitly favors <f>, with some exceptions[184] (cf. Duden 2017: 15). This adds a layer of complexity, as orthographically licensed variation is further subjected to – even if only non-binding – regulation in the form of recommendations. Notably, not every orthography has an officially codified rulebook, with the most prominent example being self-organizing English (cf. Berg & Aronoff 2017). In fact, English is only singly codified, in (semi) official dictionaries, which, however, do not serve as regulators of orthography but merely as observers and descriptors as they collect spellings that have gained their status as 'correct' through convention.

The second feature of an orthography is its (2) *bindingness*. It is not legally binding, which means there are no legal ramifications when individuals do not follow it. However, it is socially binding in that literate communities commonly accept and value its legitimacy and expect its members to adhere to its rules. This means that while deviations are not legally penalized, they are socially sanctioned in various ways (cf. Nerius 2007: 36f.). This shift of focus from the descriptive system of writing and its regularities to the norm and its prescriptive rules has been adamantly criticized: to underline how misled this prescriptivism is, Maas (2015; cf. also Eisenberg 2017) deconstructs the German word *Rechtschreibung*, a quasi-synonym of *Orthographie* 'orthography'. While the second part *-schreibung* amounts to 'spelling' (albeit without the normative connotations that might already be inherent in the English 'spelling'), the meaning of the first part, *Recht*, is commonly associated with the word *richtig* 'correct' instead of *Richtung* 'direction' and *recht* as in *es jemandem recht machen* 'to please someone'. By contrast, Maas

180 Cf. http://www.rechtschreibrat.com/ (March 18th, 2020).

181 Cf. http://www.bunka.go.jp/english/index.html (March 18th, 2020).

182 Cf. http://www.royin.go.th/ (March 18th, 2020).

183 Cf. https://www.korean.go.kr/front_eng/main.do (March 18th, 2020).

184 Most of the exceptions are loanwords such as <Graph> 'graph', <Graphem> 'grapheme', <Phonologie> 'phonology', and <Phantom> 'phantom' (cf. Duden 2017: 15).

(2015: 3) claims that it is indeed the latter two meanings that contribute to the meaning and function of *Rechtschreibung*. Thus, the primary function of an orthography should not be to determine what is correct and to sanction what is incorrect but instead to act as a recommendation of how one can write reasonably. Crucially, the central maxim should not be 'write by the rules' but 'write how you want to be read'. This line of criticism addresses exactly what has been mentioned above: the norm has eclipsed the system to such a degree that *to write* has gradually become equated with *to write correctly* with little to no tolerance for deviations (cf. Nerius 2007: 36; Schmidt 2018: 35).

Two additional aspects need to be mentioned. Firstly, it is interesting to note that the level of normativity that is inherent in the practice of writing is not shared by speech. The seldom-used term *orthoepy*, which stands for 'speaking (or pronouncing) correctly' (cf. Pabst-Weinschenk 2016) is associated with a type of normativity that is drastically less relevant in literate linguistic communities than orthography. Nerius (2007: 37) claims that this is due to the different prototypical features of speech and writing, the most important being that speech is commonly transient, while writing is permanent. Deviance from norms is thus often missed in (non-recorded) speech whereas non-standard spellings are commonly visible and traceable long after they have been produced. Another difference concerns the use of dialect. It frequently comes with connotations of different social categories and is interpreted as part of an individual's (linguistic) identity (cf. Milroy 1982). It is not so much perceived as a deviation from the norm rather than a social (regional, etc.) marker. Notably, dialectal speech is more common than dialectal writing, as in writing, the use of dialect appears much more restricted. Dialectal writing occurs, for example, in informal contexts (such as instant messaging), in prose, and poetry. Since one of the central sociopolitical functions of writing is to fix a supraregional standard variety of a language, dialect is more strongly marked in writing than it is in speech.

Secondly, users' expectations of an orthography also contribute to its bindingness: users expect the codified norm to tell them exactly what is deemed 'correct'. This is also the motivation behind the Duden's above-mentioned recommended spellings in cases in which more than one spelling is officially correct. Average users do not want to have this choice; they prefer clear rules they can obey (cf. Nerius 2007: 37; Sebba 2009: 44).

Sebba (2007) has investigated orthography as a social practice and offers a framework for categorizing deviations from the norm that only function precisely because orthography is perceived as binding. He establishes that orthography and non-orthography (i.e. deviance from the norm) can only work if the writing system allows for variants, in other words: if the graphematic solution space is large enough for the need for an orthography to even arise. If there is indeed variation, for example in the form of three possible spellings for a given word, it can be *licensed* (see below) in that "the conventional norms allow for a choice" (Sebba 2007: 30), or it can be *unlicensed* in that one or more variant(s) do(es) not conform to the norm. However, unlicensed variation is not random: It must occur in a way that "allows the original meaning to be conveyed, along with additional social meaning which derives from defying the conventions" (Sebba 2007: 30). This means that even unlicensed variation is restricted by conventionality, as unlicensed spellings still have "to be close enough to the norm to be recognisable to other members of the language community" (Sebba 2007: 32). If they were not, they would either not fulfill their communicative function or fail to convey the layer of additional social meaning intended by breaking the rules – or both. Consequently, deviant spellings are "not necessarily unsystematic" (Sebba 2007: 46). Reconsider the above-mentioned examples <Typographie> and <Typografie> for the German word for 'typography'. Add to those *<typographie> and *<typografie>, which are also found in the graphematic solution space (among many other conceivable variants such as *<Tüpografie>). The first two spellings are licensed, while the latter two – as the asterisks indicate – are unlicensed since nouns, in German, must be capitalized. Note that a reader would still be able to decode the meaning of *<typographie>, but it is the writer's choice (if it is not an unconscious mistake) of not capitalizing it that carries additional social meaning. This underlines what has been established above: spellings can be located outside of the orthographic norm but still conform to the regularities of the writing system. This is where the potential for orthography to serve as social action resides: inside the so-called "zone of social meaning" (Sebba 2007: 34), which is necessarily located outside of the orthographic norm[185] and functions on the basis that "choices are made in particular social, historical and cultural contexts" (Sebba 2007: 26). To further specify this zone of social meaning including the associated degrees of freedom to deviate as well as possible sanctions, Sebba (2007: 43f.) describes a number of regulated spaces that differ in the strictness of their regulation. In *fully regulated spaces* (such as schools, publishing houses, etc.), which are positioned at the center of the orthographic space, deviations from orthography are strictly sanctioned, while in *partially regulated spaces*, acceptance for breaking the rules is higher. Finally, in *unregulated spaces*, acceptance for unlicensed variation is highest. The boundaries between these spaces are fuzzy (cf. Figure 24).

With respect to variation, what generally characterizes orthographies is their (3) *small degree of variability*. Even if the graphematic solution space of a given writing system is fairly large, orthographically licensed variation is commonly limited to a minimum. This is intricately linked to the above-mentioned expectations that users have of an orthography, the majority of which prefer not to have to choose from a number of correct variants. Thus, what an orthography effectively does – and what lies at the core of the major criticisms voiced against it (cf. Maas 2015) – is to restrict the resources at the disposal of a writing system's users, ultimately depriving them of the possibility to choose from available resources. It hinges on the perspective, then, if orthography is interpreted as a useful regulation which prevents the possible 'chaos' that could ensue if written communication

185 Note that adhering to the orthographic norm is also a choice and, therefore, also carries social meaning.

was unstandardized or instead as a restriction that curtails the linguistic freedom of writers. It appears that commonly, the regulatory communicative function of orthographic norms is deemed much more important than any creative freedom that might be restricted by them. The line of thinking behind this is that in order to be communicatively successful, one must be understood, and unambiguously correct (because codified) spellings seem to be interpreted as a useful instrument to achieve exactly that. Variants, on the other hand, are often perceived as inexpedient and disruptive, even if they do not interfere with the comprehensibility of a written utterance (cf. Nerius 2007: 39). Comprehensibility, thus, fades into the background, while striving for uniformity through conformity to the norm comes to the forefront.

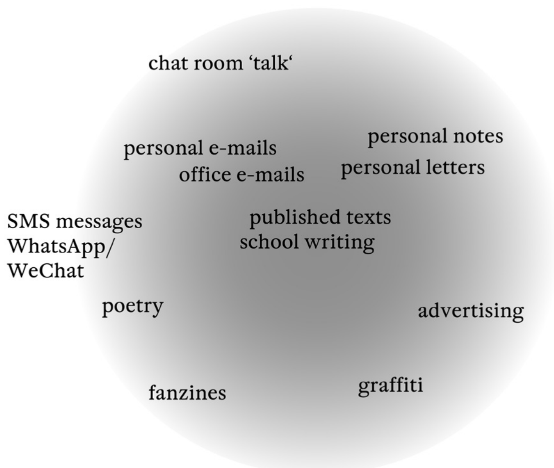



Figure 24. Orthographic space(s), from: Sebba (2007: 43)

Notably, one area in which the small variability of orthographies does filter through is the (gradual) integration of foreign material, which, as will be shown below, is a crucial task that any orthography – regardless of which (type of) writing system it regulates – has to take on. For example, an older survey of German dictionaries shows that while for less than 1% of spellings in dictionaries there exist licensed variants, 80% of them concern the spelling of loan words (cf. Gabler 1983). The question of how much variation there is in orthographies of typologically different writing systems will be discussed in greater detail below.

The last important feature or orthographies listed by Nerius (2007: 39f.) is (4) *changeability*. This 'changeability' of an orthography stands in stark contrast to the changeability of writing systems, and indeed, a more fitting name for this feature of orthographies would actually be *stability* or *rigidity*. A writing system, as described above, is rather dynamic in that through its use, it is in a state of constant change. For example, it expands when initially unsystematic forms gradually become part of the system. The interrelations between system and use are lively and natural in that they are not *a priori* regulated by norms – at least not external, codified norms. The politics behind an orthography, on the other hand, render it static. It cannot react to the actual use in the same way that the system can, as changes of the norm can only be achieved through changes of the codification. This is a task not undertaken by the language users themselves but, as a central matter of language policy, by the above-mentioned regulators, i.e. institutions of authority. That is, changes of the norm can be achieved exclusively through orthographic reforms, which are not only motivated linguistically but (one might argue predominantly) politically. Not every little change of the codification can automatically be regarded as a full-fledged reform, however, as the term 'reform' is appropriate only when general rules, i.e. rules that are generalizable over numerous contexts, are affected by the change (cf. Nerius 2007: 40 and the next section). If instead, only isolated cases – single spellings, for example – are changed, these minor modifications do not constitute a reform. However, even in the most minor case, the codification effectively has to be changed. Also, the fact that norms make it possible for users to consult the correct spelling at any time contributes greatly to the stability of orthography (cf. Kohrt 1990: 116).

## 3.2 Types of orthographic rules

The most central linguistic phenomenon relevant to orthographies is the concept of *rule*. Unlike the descriptive rules (to be read as 'regularities') of other linguistic domains – such as syntax, morphology, phonology, but also graphematics – orthographic rules are of prescriptive nature. They are to be interpreted as instructions to produce spellings that conform to the norm (cf. Ewald 2007: 43). At this point, I want to mention explicitly the conceptual difference between rules and explications of rules. As Kohrt (1990: 108, my translation) notes, "the same entity can never be part of a linguistic norm and simultaneously serve as the description of that exact norm". In other words, the object level and the meta-level must not be conflated. To distinguish between these two phenomena, I refer to the former as *rule* (or *norm*) and to the latter as *rule explication* (or *norm explication*).

Several subtypes of orthographic rules must be distinguished: Firstly, *given rules* must be separated from *set rules*. The difference between them can be explained by considering the aforementioned relationship between system and norm. *Given rules* correspond with graphematic regularities, they are rules that have arisen organically through the use of the system. The emergence of such rules can be explained by reference to *invisible hand theory* (cf. Keller 2014): language users, in this case acting as writers, have implicitly agreed on conventions which potentially develop into rules. The attribute *given*, thus, underlines that these rules are already *given* in the system, and their rule explications merely codify them. In this context, Ewald (2007: 42) argues that when given rules are codified, they are externalized through this codification and, thus, also become *set rules* (see below). This might be accurate, but for the sake of conceptual and terminological clarity, I will still speak of *given rules*, highlighting that they are codified orthographic rules that are nonetheless based in the system and its use. This type of rule corresponds with what Neef (2015) calls *systematic orthography*. After *given rules* have been transferred from graphematics to orthography, their application to the graphematic solution space might leave more than one possible candidate for an orthographically correct spelling, such as the possible spellings <write> and <right> for *write*, which both conform to the graphematics of the English writing system. In such cases, "[c]onventional orthography [...] decides which of these options the correct one is" (Neef 2015: 720). All of the variants that *conventional orthography* may choose from – such as the mentioned <write> and <right> for *write* – obey given rules and are part of the system. Thus, the process of choosing one of them as the correct spelling for a given word is either arbitrary or determined by the second type of rules, *set rules*.

*Set rules* are rules which exist only in their externalized form. They "are dependent on a metalinguistic objectification and only function because of being explicitly recorded" (Ewald 2007: 42, my translation). In other words, they are rules that are dependent on their rule explication; even more so – they are constituted by their rule explication. Unlike given rules, which correspond with the regularities of the system and thus mostly with users' internal norms, they are commonly obeyed consciously by users. In Figure 23 above, these rules fall, at least right after they are constituted by a rule explication, into that part of the norm that shows no overlap with the system. Notably, they can still overlap with the use in cases in which they react to the use and codify unsystematic but nonetheless used spellings. They can also overlap with the use when users react to novel codifications (e.g. right after a new reform has been implemented) by starting to use correct but (yet) unsystematic spellings. Finally, it is also possible that set rules fall into the area of the norm that overlaps neither with system nor use. This situation arises when regulators – driven by different motivations – construct and codify unsystematic spellings that are not yet in use. In such cases in which spellings are codified as orthographically correct even though they are not part of the graphematic solution space, I would speak of *unsystematic orthography* as opposed to *systematic orthography*. Unlike systematic orthography, unsystematic orthography is necessarily always conventional. This explains why Neef (2015: 720) posits that "conventional orthography cannot be fully reconstructed as a theoretical system, [...] only partially". It is the systematic part than can be reconstructed, while the unsystematic part cannot.

An example comes from German, where spellings of loan words have been codified that were not grounded in actual use. But the Council for German Orthography had to realize that even after they were codified, they never came into use, which led to some of them being dropped from dictionaries. For example, in its 27th edition, the Duden reacted to the 2016 report of the Council and dropped spellings such as <Ketschup> 'ketchup', <Joga> 'yoga', <Grislibär> 'grizzly bear', and <Majonäse> 'mayonnaise' because they are not (anymore) in use (cf. Duden 2017: 18).

With respect to the scope of rules, three categories can be established: *principles*, *general rules*, and *singular rules*. The first term, *principle*, is very prominent in the German-language literature on orthography, although it is not unproblematic as it contributes to blurring the line between graphematics (system and use) and orthography (norm). When principles are listed, one can commonly find the phonological (alternatively also phonetic, phonographic), syllabic, morphological, economic, lexicosemantic, etymological, aesthetic, grammatical, textual, and pragmatic principles, among others (cf. Karg 2015: 48–69). To introduce yet another distinction, following Rahnenführer (1989), *principles of writing*, which are graphematic in nature, need to be kept apart from *principles of orthography*, which are, as the name suggests, orthographic. Here, admittedly, a neat separation of graphematics and orthography is really challenging. When Kohrt (1990: 112, my translation) mentions 'principles', for example, it is unclear which of the two he refers to:

> […] 'principles', which are essentially valid cross-linguistically and which, in particular contexts, and more or less strongly, prevail in different stages of the development of a language in what was effectively a 'free play of forces.'[186]

Elsewhere, he observes that principles are

> essentially an expression of an *a posteriori systematization* of what language users participating in writing in alphabetic writing systems have done, and use of [the term 'principle'] rests basically on the mere *insinuation* of action-driving maxims that determine the doing of individuals in a given practical area.
> (Kohrt 1987: 516, my translation, emphasis in original)[187]

I argue that what Kohrt describes are *principles of writing*, possibilities of encoding different facets of linguistic information – most importantly sound vs. meaning – in various ways. This already anticipates the discussion that principles are in conflict with each other – a situation Kohrt refers to as a 'play of forces'. If writers can choose to spell the noun deriving from German <kalt> 'cold' either as <Kälte> or as <Kelte>, they have to decide what is more important to them: phonographic transparency (the phonological/phonetic/phonographic, etc. principle) and with it a basic shape-phoneme correspondence that is closer to biuniqueness, or instead morphographic transparency (the morphological principle), where, however, the one basic shape–one phoneme ideal is sacrificed in the process. Thus, they can choose from different resources and possibilities that the graphematic module of a writing system offers. At this point, no orthographic norms or rules are intervening yet. However, these general principles of writing can be operationalized and transferred to a normative context. This is precisely what Kohrt means when he

186 „[…] ‚Prinzipien, die grundsätzlich einzelsprachübergreifend gültig sind und sich in den verschiedenen Stadien der Entwicklung einer Sprache in einem quasi ‚freien Spiel der Kräfte' an einzelnen Stellen jeweils mehr oder minder stark durchsetzen."

187 „[…] wesentlich Ausdruck einer *nachträglichen Systematisierung* dessen, was die Sprachteilhaber beim Schreiben mittels Alphabetschrift getan haben, und seine Verwendung beruht im Grunde auf einer bloßen *Unterstellung* von handlungsleitenden Maximen, die das Tun der Individuen in einem bestimmten praktischen Bereich bestimmen."

speaks of an *a posteriori systematization*: Orthographies, at least in their initial stages, draw heavily from the system and its use, and on the highest level of abstraction, this concerns the general possibilities of a given writing system to refer to the different levels of language – phonology, morphology, lexicon, etc., and later include also possibilities that are not necessarily driven by structural linguistic reasons, cf. the aesthetic or the pragmatic principles. After having been transferred from the system to the norm, these principles of writing indeed become principles of orthography (or *orthographic principles*). They consequently serve as 'action-driving maxims' and can be instrumentalized by language regulators as a template in the design of new rules.

This sheds light on why it is so hard to draw a line between principles of writing and orthographic principles. In sum, principles of writing are ontogenetically primary, and the secondary orthographic principles can only draw on what the principles of writing are offering. For this reason, orthographic principles, very general in nature, cannot be changed the same way that orthographic rules can (cf. Rahnenführer 1989: 291). They are not entirely without normative force, but at the same time, they are not as normative as rules as they lack the role of instructions (cf. Naumann 1990: 149; Rahnenführer 1989: 290). Instead, they provide orientation, which is why the suggested term *orthographic orientation* is probably more fitting (cf. Kohrt 1987: 509f.). Since they offer some degree of normative 'guidance' and serve as the basis for general rules, they are sometimes also interpreted as rules of a higher order, so-called *hyperrules* (cf. Kohrt 1990: 106).

The cross-linguistic applicability of principles of writing and orthographic principles must be commented on. Principles of writing, as described above, represent general possibilities of how the graphic units of a writing system can represent different levels of language. These principles are not specific to a given writing system but are characteristic of types of writing systems, which Kohrt implies when he singles out 'alphabetic writing systems'. It is the general nature of possible graphematic relations that determines the linguistic levels that can be represented by a writing system. In other words, the principles represent the possibilities of a given type of writing systems in that they are not specific to individual systems but can also not be generalized across types of writing systems. In fact, types of writing systems represent exactly an *a priori* prioritization of one principle of writing, which leads to the broadest categorical distinction in writing system typology: phonographic writing systems, which rely heavily on the phonological principle (and related principles such as the syllabic principle), and morphographic writing systems, which are grounded in the morphological principle.

Unlike Rahnenführer (1989: 288), I do not believe that the orthographic principles derived from principles of writing are specific to given writing systems. Notably, the relevance and relative hierarchical ordering of orthographic principles can differ even in orthographies of writing systems that are assigned to the same type: while some alphabetic orthographies favor the phonological principle (e.g. Finnish), other orthographies emphasize the morphological or 'etymological' principle (e.g. English). The repertoire of general principles of writing that orthographies can draw on, however, is always the same, which means that orthographic principles are not inherently writing system-specific. What is system-specific, however, is *orthographic rules* for which orthographic principles, as hyperrules, serve as a basis.

With respect to orthographic rules, *general rules* and *singular rules* must be distinguished. In a nutshell, general rules do not affect single spellings but groups of spellings, whereas the scope of singular rules is limited; they affect only isolated cases (cf. Ewald 2007: 44–48). Thus, general rules, as subrules of orthographic principles, have a larger scope and apply in more cases, while singular rules work in the smallest possible context: a single word, or maximally a few words. Traditionally, singular rules are interpreted as exceptions: either because they cannot be subsumed under a general rule or because they contradict a general rule, which is the case when a general rule fails to apply in a context in which it usually applies. An example of a general rule explication is the codification of the use of <ß> in some varieties of German: when /s/ occurs after a long vowel or a diphthong and is not followed by a consonant in the stem of the word, <ß> is generally written (cf. Amtliche Regelung 2018: 29, § 25; Ewald 2007: 45). Note that this general rule explication states a context and gives instructions on what must be done in that context. By contrast, singular rules are of special relevance in contexts in which we encounter so-called 'false' general rule explications, such as § 18 in the codification of German orthography: "In few words, the diphthong [aɪ], as an exception, is spelled *ai*" (Amtliche Regelung 2018: 24, § 18, emphasis in original, my translation), with an addendum stating "[t]his concerns words such as *Hai* ['shark', D.M.], *Kaiser* ['emperor', D.M.], *Mai* ['May', D.M.]".[188] Since the provided list of exceptions spelled with <ai> is not exhaustive, this supposedly 'general' rule explication including its addendum does not suffice to correctly spell /ai/ in all contexts (cf. Gallmann & Sitta 1997: 95). This is where singular rules are needed: in order to know how to correctly spell words containing /ai/ that are not mentioned in this short list, a writer must look them up in a dictionary.

This is also where the above-mentioned *double codification* of orthography comes into play: general rule explications such as the scope of <ß> cited above are codified in sets of orthographic rules. Singular rule explications, on the other hand, are codified in (orthographic)[189] dictionaries in which every entry represents a singular rule explication. Initially, orthographic dictionaries resem-

**188** „In wenigen Wörtern schreibt man den Diphthong [aɪ] ausnahmsweise *ai*. [...] Das betrifft Wörter wie: *Hai, Kaiser, Mai*."

**189** Kohrt (1990: 119) comments on the fact that lexicographers usually – with few exceptions – use orthographically correct lemmas in all kinds of dictionaries: this, according to him, makes it difficult to distinguish true 'orthographic' dictionaries from other kinds of dictionaries which are, given that they offer exclusively orthographically correct spellings, also perceived by users as orthographic dictionaries, irrespective of their intended function or use (e.g. an etymological dictionary). Thus, 'true' orthographic dictionaries are those (monolingual) dictionaries in which it was the lexicographers' intent to allow users to look up the correct spelling of a word.

bled appendices in that they offered illustrative examples of the rules that were laid out in the preceding sets of orthographic rules and demonstrated how the rules were to be applied in practice (cf. Schaeder 1986: 199). Interestingly, Kohrt (1990: 110) observes that there appears to be an inverse relationship between the public's interest and the theoretical interest in dictionaries: while the addition (and, thus, codification) of new words to dictionaries, the change of existing spellings, etc. is often a matter of public debate, grapholinguistics and other disciplines have neglected the relevance of dictionaries in the codification and negotiation of orthography.

A final interesting issue with respect to rules and their codification must be addressed: codifications are open-ended, which renders orthographies *open normative systems* or *open standardizations* (cf. Kohrt 1990: 129). This issue also concerns the question of when exactly a word can be regarded as a word, with the public often adhering to the belief that a word is only a word when it is listed in the dictionary.[190] This, of course, represents prescriptive thinking and is usually rejected by linguistics, as the existence of words and their status as words do not depend on them being included in a dictionary. However, if words are not included in the dictionary, the question of their orthographically correct spellings must remain unanswered – they are orthographically undetermined.

Up until this point, we have dealt exclusively with *absolute norms*. They are absolute in that they can either be obeyed or not: if they are obeyed, a written utterance will be orthographically correct; if they are not obeyed, it will be orthographically incorrect. This distinction is categorical, there is no 'more or less correct' in between. However, I argue that this is not the only kind of norm and that there exist also *loose norms* in the form of recommendations that should generally be followed but whose rule explications are not explicit enough to always allow a judgment of whether a spelling is orthographically correct or not. The distinction between absolute norms and loose norms corresponds with what Naumann (1990) terms *orthography* and *graphostylistics*, respectively, with orthography prescribing strict norms and graphostylistics offering recommendations.[191] To illustrate this, I want to discuss two examples, one from German and one from Chinese: In German, not unlike in English, to mark a parenthesis, a writer can use either a pair of complement opening and closing parentheses <( )>, two dashes <–>, or two commas <,> (but cf. Bredel 2011: 144f. for pragmatic restrictions on the

---

**190** Consider, for example, the rules of the popular board game *Scrabble*. Player A might lay down a word and player B might 'challenge' it by claiming that it is not a 'real' word. In this case, the rules state that players must consult a dictionary (one they have formerly all agreed on) to check if the word is included. Only if it is included (and conforms to an additional number of rules: not being an abbreviation, not always being capitalized, not including hyphens or apostrophes, not being an affix that stands alone) does it count as a 'word' (cf. Scrabble Rules n. d., https://scrabble.hasbro.com/en-us/rules, March 18th, 2020).

**191** Note that the term *graphostylistics* has also been used in a different sense, for example by Spillner (1974), who used it to refer to graphetic style choices.

use of the latter two). Which of them is chosen is not orthographically determined, as all of them are orthographically licensed for marking parentheses. Since a writer can freely choose between them, they can be considered (almost) equifunctional orthographic variants. Through individual choices, writers have the freedom to act creatively, which opens up the space of social meaning, making Naumann's term *graphostylistics* fitting for this type of situation. Note, however, that the same does not apply to the same degree to orthographically licensed variants such as <Typographie> vs. <Typografie>. In this case, writers are expected to make a definite intra-textual choice: once they have chosen <Typographie>, they cannot[192] alternate it freely with <Typografie> in the same text. (An inter-textual alternation is, of course, possible, although it is expected that most writers will idiolectally stick to one variant.) In comparison, the three different kinds of demarcations for textual parentheses can be alternated freely within the same text, and it is this complete choice of freedom that corresponds with the traditional meaning of *stylistics*.

The situation in the Chinese example is a bit different: a frequently mentioned challenge for the Chinese writing system is the integration of foreign names or words. In theory, graphemes can be used solely for their phonographic value while their morphographic value is being ignored. This is commonly practiced when foreign names are transcribed, a process in which the phonological level is foregrounded and morphological information becomes secondary. The challenge, now, arises due to the relatively small number of phonotactically licensed syllables in Mandarin and the resulting multitude of homophones. Since the Chinese writing system is morphographic, morphemes – many of which are homophonous – are written with distinct graphemes. In Modern Chinese, no new characters are created for writing foreign names, "probably for typographic reasons and/or due to the lack of familiarity of newly created characters" (Hsieh 2015). Instead, existing characters are used and desemanticized. There are certain guidelines 'to a good translation', i.e. an adequate choice of characters for a given name, which are *faithfulness*, *expressiveness*, and *elegance*. Corpus studies have shown that these guidelines are followed by Chinese writers, although it remains unclear whether this happens consciously or unconsciously (cf. Hsieh 2015). What the practice of transcribing names reveals is that the phonetic similarity of characters to the foreign name to be written is often sacrificed for semantic considerations. For example, it is recommended that no characters with negative connotations be used. Also, foreign names are to be somehow marked to indicate their 'foreignness'. This can be – and is frequently – done by choosing low-frequency characters. However, there now exist prescriptive regulations by the Xīnhua News Agency which discourage the use of low-frequent characters. Additional "prescriptive norms [...] require that characters with too many strokes or heteronyms should not be used"

192 'Can' is likely not the fitting verb for this situation. Of course, writers *can* do it, but ironically, it would probably be regarded as breaking the rules. As both spellings are orthographically correct, writers have not made an actual mistake. However, consistency of spelling is also expected from writers, which, arguably, also makes it a part of orthography.

(Hsieh 2015). Unfortunately, it is not straightforwardly clear just how binding the regulations published by the mentioned news agency truly are as well as whom they are directed at, and it is vague which norms Hsieh is exactly referring to in the last quote. However, it is clear that we are dealing with a form of *loose norms* and *graphostylistics* here. With only loose norms, there exist neither general rules nor singular rules for writing a given syllable (or, more directly, morpheme). The two syllables /di/ and /mi/ of my nickname Dimi, for example, can be written with a variety of homophonous graphemes. The existing norms that no characters with negative connotations – the perception and evaluation of which is arguably subjective – and no rare characters be used reduce the number of possibilities, but they do not determine one and only one correct choice. In other words, the general rules are too general, and there exist no singular rules since foreign names are not included as entries in dictionaries. In sum, in cases like these, users are confronted with freedom of choice, and the loose norms in question leave open a number of variants that all are to be considered as orthographically correct. They might not be equally correct (or better: appropriate), however, as in the Chinese example, where 'correctness' is not an absolute but a scalar category, some spellings for /di/ are perceived as more appropriate for writing foreign names than others.

## 3.3 Sketch of a comparative orthography

Having turned to Chinese, I want to point out the problematic fact that most – if not all – of the preceding examples in this chapter have come from German, reflecting the fact that most contributions to orthographic theory have come from German-language literature and have also focused almost exclusively on German. Unfortunately, thus, Eurocentrism has not spared research on the module of *orthography*, either. One of the likely reasons for this has already been addressed: in the Anglo-American realm, and, consequently, in most of the relevant grapholinguistic literature published in English, *orthography* is understood as having a descriptive meaning and being used as a synonym of *writing system*.[193] This is not inherently erroneous, but it is misguidedly reductive in descriptive linguistics: *orthography* should designate exclusively the standardization (or, from the point of view of its product, the standardized versions) of writing systems. The result of equating *orthographies* with *writing systems* is 'descriptions' that are caged in prescriptivism. These are problematic insofar as they exclude crucial and interesting phenomena that fall beyond the scope of what is *correct* but are still part of the

**193** Using *orthography* and *writing system* interchangeably within the same work – which some authors do – does not improve the situation, as it only shows that the differentiation between system and norm is disregarded, and it is up to the reader to decipher which of the two the terms ultimately refer to. In sum, using *writing system* and *orthography* as synonyms that designate the same phenomenon is problematic as it treats the actually distinct phenomena as hierarchical equals and neglects the fact that an *orthography* is actually only (an optional) part of a *writing system*.

system. As established above, not all resources available in a writing system are relevant for its orthography.

That being said, it is mostly scholars with some insight into the German grapholinguistic tradition who treat orthography as a prescriptive, normative phenomenon.[194] One of them is Florian Coulmas, who writes in his *Blackwell Encyclopedia of Writing Systems* that orthography stands for

> [c]orrect spelling and that part of grammar that deals with the rules of correct spelling. An orthography is a normative selection of the possibilities of a script for writing a particular language. All orthographies are language specific. As the most visible and most consciously learned linguistic subsystems, orthographies are often codified by official decree. (Coulmas 1996a: 379)

Two passages in the introduction to this encyclopedia entry are baffling: that orthography is a 'part of grammar' and that it is a 'consciously learned linguistic [subsystem]'. While all the other mentioned aspects – the focus on correct spelling, the normative selection, the language specificity, the codification by official decree – correspond with what has been described above, interpreting orthography as a part of grammar is arguably problematic. Since the concept of orthography can only be understood in combination with the concepts of *writing system* and *graphematics*, it is necessary to also look at these respective entries to analyze in what relation they stand to each other. What is immediately obvious from reading the entry on *graphemics* (as an alternative designation for *graphematics*) is that it serves solely as the designation for the "linguistic study of writing systems" (Coulmas 1996a: 176), meaning it is only used for the meta-level and not interpreted as also designating a subsystem of writing systems, thus leaving this spot open. Coulmas fills this open spot with one reading of *writing system*: "a writing system is what is also referred to as spelling, i.e. a system of rules underlying the use of the graphemes of a language" (Coulmas 1996a: 560). Ultimately, the only explicit difference between *writing system* and *orthography* in Coulmas' entries is that the former deals with 'rules of spelling' and the latter with 'rules of correct spelling'. Only the first is part of the grammar of a language system equipped with a writing system, the second is externally codified. (However, it certainly can, as mentioned above, be internalized and become part of the grammar).

An interesting section of Coulmas' entry on *orthography* deals with the crucial question of *what* can be standardized and regulated in different writing systems. Herein, as I argue below, lies the crux of the Eurocentrist treatment of orthography which, quite simply, is primarily a reflection of the lack of research on orthography in general. Coulmas (1996a: 379) observes that "[i]n alphabetically written languages, the aspects of writing most commonly codified by means of orthographic rules are grapheme-phoneme correspondence, word division, hy-

194 However, German authors are also not without fault. For example, a prominent German textbook (Fuhrhop 2015) that offers a description of the graphematics of German is inadequately entitled *Orthografie* (cf. Dürscheid 2016: 164) when it should of course be *Graphematik*.

phenation, capitalization, and the spelling of loan words". He goes on to note that "sound-letter correspondence is also a central component in orthographies of other writing systems" (Coulmas 1996a: 380), positing that orthographies *of* non-alphabetic writing systems[195] share properties with orthographies of alphabets. He continues by adding that indicating vowels is orthographically relevant in Hebrew and Arabic and that in Chinese, "the graphic composition of characters is a matter of orthographic regulation". While Coulmas does mention that in different writing systems, different features are orthographically regulated, the examples from non-alphabetic systems remain sparse. Notably, most of the aspects that are relevant in alphabetic writing systems have no correlate in non-alphabetic systems: capitalization is not present in any other type of writing system except for alphabets, while word division is only present in some (but not Chinese, Japanese, or Thai, for example), and grapheme-phoneme correspondences are also relevant solely in segmental phonographic systems (in syllabographic systems, graphemes do not correspond with phonemes, but syllables; in Chinese, they do not directly correspond with any phonological unit at all). Thus, the listed aspects are not or only marginally relevant for a broader picture of orthographic regulation.

Twenty years after his encyclopedia entry, Coulmas (2016: 41) notes that "[…] it is a legitimate question whether the structural differences between them [= writing systems, D. M.] have any implications for prescriptive rules and attitudes". With this question, he captures what comparative orthography research needs to investigate: what is regulated in different writing systems and how can differences in the focus of regulation be explained? Are the differences of linguistic nature, that is, rooted in the writing systems that are regulated? Or are they politically or culturally determined?

Other central questions of comparative orthography research are: Who is in charge of the standardization of a given writing system? What are the motivations behind standardization efforts? In which parts of a given writing system does variation occur and are they the focus of regulation? Put more simply: which aspects in a given writing system can be and/or are prone to being standardized, and are all of them actually standardized or are some of them attended to more than others? If the graphematic solution space offers variants, how does an orthography (or, less anthropomorphically, the stakeholders deciding on it) choose one (or multiple) of them? How can an orthography intervene to 'fix' or compensate features of writing systems that are unsystematic and/or complicate its use? How do different orthographies respond to the system and its use? Are they grounded in actual use? What ultimately drives orthographies? Is it a linguistic analysis of the writing system? Politics? Another question is concerned with the ontogenesis of orthography: what was the motivation behind the development/implementation of a given orthography? What are the historical circumstances in which it began to emerge?

195 This passage strongly implies Coulmas interprets orthographies as parts of writing systems, much as I do in this book.

## 3.4 Orthography reforms

As the preceding treatment of orthography as well as the list of questions just mentioned imply, there are several concurrent discourses that address orthography: as Eira (1998) showed in her study of authority and discourse, these are not limited to scientific discourses but include also political discourses and religious discourses (see Figure 25). Schimmel-Fijalkowytsch (2018) adds discourses negotiated by the media. Ultimately, as a central concern of language policy, orthography is both a linguistic and a political matter. Thus, several agents or 'stakeholders' take an interest in it and want to be involved in shaping it: academics, government representatives, but also religious institutions such as missions or priesthoods.

A crucial type of discourse that is not sufficiently captured in the literature but must not be forgotten is public discourse. When people reflect on language in everyday life, it is often actually writing that they think about, and arguably, specifically orthography, since their considerations are frequently of prescriptive nature: 'How do I write/spell this or that *correctly*?' is one of the most common language-related questions literate people ask themselves. Orthography is much more palpable than other linguistic domains such as phonology or morphology. The material permanence of writing allows and sometimes even invites metapragmatic reflection on writing and orthography. Also, orthography is a pressing issue for the public given that it is commonly a main concern in primary education and most people in literate communities come into intensive contact with it in the course of literacy instruction. Other linguistic domains are usually not treated this consciously – unless one studies linguistics (or related subjects treating linguistic issues). In sum, the public often interprets orthography as *pars pro toto* for language, and many people feel they should have a say in what happens with their language.

When changes in the orthographic codification are discussed, not only the public but many stakeholders with different motivations want to be included in the discussion. As Eira (1998: 175) argues, academics commonly have the linguistic fit of an orthography in mind (i.e. how well it suits the language for which it is used), while politicians and religious authorities focus on other goals, respectively. Interests of linguists and politicians intersect with respect to education: many orthographic reforms were put in place to facilitate the acquisition of reading and writing with the goal of increasing literacy rates (for the example of the simplification of Chinese characters for precisely this reason, cf. Handel 2013; Section 5.2). Even though the intended goal is the same, politicians and linguists want to achieve it for different reasons: politically, an increase in literacy rates might contribute to higher economic competitiveness, while linguistically, the motivations behind achieving more widespread literacy might appear more altruistic and idealistic. However, it is important to note that linguists are aware of and want to underline their expert knowledge concerning orthography and want to be seen and respected as integral (and authoritative) parts of any language-related decisions.

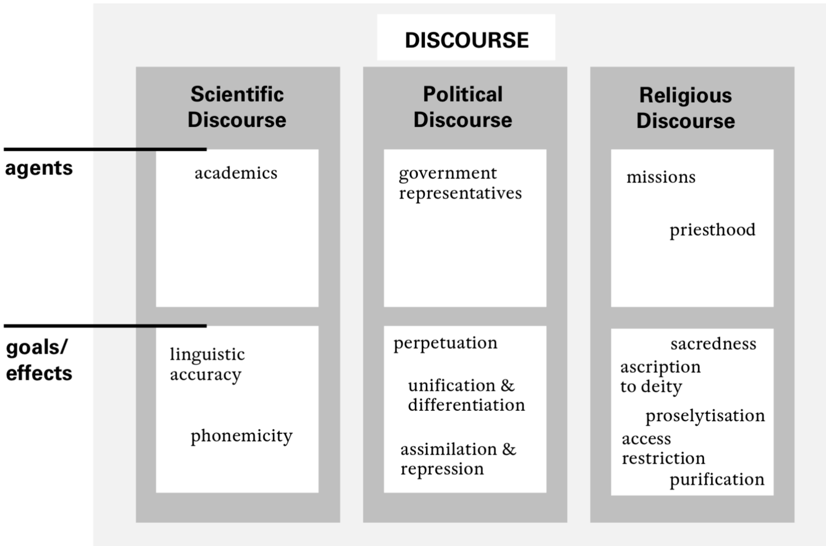



FIGURE 25. Discourses surrounding orthography, from Eira (1998: 175)

Given that numerous different stakeholders are invested, orthography reforms often evoke considerable debate. The reform of German orthography in 1996 is a striking example as it sparked a number of protests and even resulted in private citizens constitutionally challenging the reform (for detailed accounts, cf. Johnson 2002, 2005; Schimmel-Fijalkowytsch 2018). Again, the motivations behind this outcry are multifaceted: while the public mostly wants to adhere to the version of an orthography they have grown accustomed to, linguists and professionals working in education might protest suggested changes if they are incongruent with the underlying system and, in consequence, might not lead to a facilitation of reading and writing.

This is not the place to discuss the different discourses surrounding orthography in detail. However, whenever orthography is being addressed, it should be kept in mind that it is not only a linguistic matter but also a deeply social and, thus, sociolinguistic matter (cf. Unseth 2005; Sebba 2009). When seen as a topic of sociolinguistics, the different facets of orthography (linguistic, ideological, political, religious, etc.) actually merge. The resources of orthography have myriad functions and have the potential to convey meaning on various levels beyond the denotative. Orthography is not only an optional module of a writing system but, as the surface realization of a writing system, the central and phenomenologically primary interface between a system and its users. While this book's focus is on its status as a module of writing systems and, thus, its linguistic functions, the fact that orthography represents "social action" (Jaffe et al. 2012) shall at no point be undermined. In fact, in the discussion of the sociocultural fit (cf. Chapter 8), this perspective will come to the forefront.

# III Explanation

In this book, a distinction is being made between description and explanation and, consequently, between descriptive theories and explanatory theories. What must be noted in advance is that while 'explanatory', as a term, is often perceived as being intimately tied to the generativist paradigm, it is not used as such here since the aim is to sketch a functionalist theory of writing. In short, and very simplistically, with respect to grapholinguistics, description deals with the question of how writing systems are built and how they are used, whereas explanation strives to answer why they are built the way they are and why they are used as they are. Crucially, two pairs of questions are intimately tied to one another, with one being located within the scope of description and the other being attended to by explanation, respectively: how writing systems are used affects why they are structured the way they are, and how writing systems are structured has a bearing on why they are used the way they are. This shows that description and explanation (should) go hand in hand in a theory that is neither merely descriptive nor just explanatory. Notably, the relation between description and explanation is not symmetric: the 'why' questions can be answered only when description has already been attended to, while the descriptive 'how' questions, as the very basis of (grapho)linguistic research, are not dependent on explanation. In other words, description precedes explanation, which motivates the structure of this book, and moreover, description is obligatory, explanation is optional. The remaining question, now, is why is explanation worth striving for?

A fact that is so obvious that we do not even need description to become aware of it – as mere observation suffices – is that writing systems are diverse. First of all, the myriad different scripts make them look different, and even preliminary analyses show that they function in sometimes remarkably different ways. What we need more sophisticated description for is to see that despite their differences, writing systems also share commonalities. These might be located at more abstract levels than the conspicuous differences, but they exist. Description were enough if we were ready to leave it at that. The much more interesting question, however, is where these commonalities come from, and for that, we need explanation (cf. Watt 1998: 118, specifically the quote cited in the introduction to the present book). An explanatory theory also better captures the complexity of the phenomenon that is writing by acknowledging not only how it is structured and how it relates to language, which is an important but not the only linguistically relevant aspect to writing, but also how it is used by humans – both from psycholinguistic and sociolinguistic perspectives. This is in no way meant to devalue descriptive theories (as also palpably evidenced by the extensive descriptive treatment of writing systems that precedes this part); as mentioned above, they are a necessary basis for explanatory theories.[196] In the end, it is about which questions one seeks to answer. In other words, diverse epistemological interests justify all kinds of different theories.

Explanatory theories, at least when they are functionalist, also require more than do descriptive theories. As Dryer (2006: 212f.) puts it with respect to general linguistic theories: "The nature of functionalist explanation is such that it is external to the grammar, not only in the sense that the theoretical concepts appealed to in the explanation lie outside of grammar, but also (and perhaps more controversially) because there is no way to build these explanations into the grammar itself." What is needed, thus, is not only a diverse (grapho)linguistic data sample but also extra(grapho)linguistic evidence, especially from use. This is at the core of usage-based approaches to linguistics, and also of Naturalness Theory (cf. Section 4.2), the functionalist approach that will be used here as a starting point of a theory of writing. The fact that in this part of the book, references will be made to evidence stemming not only from linguistics but a strikingly broad range of disciplines (including psychology, the cognitive sciences, sociology, anthropology, information science, design theory, paleography, etc.) also fits and aptly underlines the interdisciplinarity of grapholinguistics. The fact that writing is not a subject that can be comprehensively studied by a single discipline and that explanation – more so than description – brings together the diverse relevant perspectives also strengthens the decision to pursue explanation.

One way to get to explanation is through the angle of evaluation, as it implies the question of why something is 'good' or 'bad' for a given purpose. Strikingly, as will be discussed in Section 4.1, writing systems have in the past frequently been evaluated – and sometimes compared – with respect to their 'goodness'. These past evaluations – even if they are often unsystematic – are noteworthy for two reasons: firstly, they mention criteria that could be useful in explaining the structure and use of writing systems (or, more specifically, their three modules of graphetics, graphematics, and orthography), and secondly, they can be assigned to four categories that strive to answer the following questions, respectively: 1) How systematic is a writing system?, 2) How well does a writing system relate to its language?, 3) How well is a writing system suited for physiological and cognitive processing by humans?, and 4) How well is a writing system suited for communication and to fulfill users' sociocultural wishes and needs? These central questions

196 Cf. also Dryer (2006: 213): "A grammatical description of a language is thus not deficient or inadequate if it leaves out explanations for why the language is the way it is. In fact, in so far as grammars exist independently of explanation, there is a need for description independent of explanation."

are captured by so-called 'fits' – the systematic, linguistic, processing, and socio-cultural fits – which lend this part of the book its structure.

# 4 Prolegomena

## 4.1 Evaluation of writing systems[197]

> I have taken it as given that some writing systems are better than others. This could certainly be debated. (Rogers 1995: 31)

The idea that writing systems can be evaluated is not new. Indeed, many publications have painted a picture of an optimal writing system based on a variety of criteria that writing systems must fulfill and/or features they must exhibit. This practice becomes most evident in guidelines that describe how new writing systems should be devised: in *Advances in the creation and revision of writing systems* (Fishman 1977), for example, a contribution is titled *Principles for the design of practical writing systems* (Venezky 1977), reflecting that writing systems are qualitatively evaluated based on criteria, or, in this case, "principles". Titles of other works that are ostentatiously indicative of this are *The ideal orthography* (Bauernschmidt 1980), *Factors in designing effective orthographies for unwritten languages* (Cahill & Karan 2008), *In search of the perfect orthography* (Venezky 2004), or *Optimal orthographies* (Rogers 1995). What, now, makes writing systems practical, ideal, effective, perfect, or optimal? Many of the criteria discussed in works such as these were, for the most part, postulated intuitively. There is nothing inherently wrong with such an inductive approach, and the criteria that it results in are by no means irrelevant. However, in an attempt to systematically construct a functional grapholinguistic theory, criteria should not simply be appointed without an explanation of their underlying foundations as well as (external) evidence supporting them.

Aside from the mentioned evaluative treatments that often aim at optimizing the development of new writing systems, there is also mere and blatant ethnocentrism, predominantly Eurocentrism, which basically culminates in the view that the alphabet is the most ingenious invention of Western civilization (cf. the *Alphabet Effect*, Logan 2004, and criticism thereof in Grosswiler 2004; cf. also Olson 1996: 8f.). According to this view, the alphabet is distinguished sharply from other types of writing systems such as syllabaries or morphographic writing systems, which is most evident in statements such as "it is generally accepted on all grounds an alphabetic system is the best" (Berry 1958: 753) or, to mention a more

197 Parts of this section have appeared in Meletis (2018).

recent claim, "[o]verall, it is argued that morphographic systems are inferior to phonographic ones" (cf. Jones & Mooney 2017: 13). This alleged "superiority of the alphabet" (Barton 1995: 20) and the simultaneous depreciation of writing systems such as Chinese and Japanese (cf., for example, Hannas 1997) have had major repercussions for linguistics and especially grapholinguistics: not only was the diachronic development of writing modeled in an evolutionary framework that propagated a teleological movement in stages starting with pictographic writing and ending with the final and optimal end-stage of alphabetic writing (Gelb's infamous and refuted *Principle of Uniform Development*, cf. Gelb 1969: 201; for criticisms, cf. Mattingly 1985; Miyamoto 2007). But this alleged superiority of the alphabet resulted also in what Share (2014) calls an "alphabetism" that pervaded and to this day largely infiltrates interdisciplinary grapholinguistic research. This includes not only the modeling of reading and writing processes but also descriptive models of writing and grapholinguistic theorizing in general.

Table 6 collects some of the criteria that have been mentioned in the grapholinguistic literature. Notably, some of them concern graphetics, others graphematics or orthography, and some apply to all of them. This, once again, highlights the fact that these notions are frequently not kept apart even in pertinent grapholinguistic literature. Most of the listed criteria will recur in the following chapters. Although at first glance, the lists appear in and of themselves diverse and unsystematic, the criteria can be assigned to four categories that each covers one perspective of how well the constituent modules of writing systems – graphetics, graphematics, and, optionally, orthography – meet certain requirements: the 1) *systematic fit*, the 2) *linguistic fit*, the 3) the *processing fit*, and 4) the *sociocultural fit*. In short, the systematic and linguistic fits are evaluated descriptively and semiotically, the processing fit is of psycholinguistic nature, and the sociocultural fit introduces a predominantly sociolinguistic perspective. Interestingly, these three fits cover what is also included in the comprehensive definition of (extra)linguistic foundation in linguistic Naturalness Theory (see below), namely physiology, semiotics, cognition, and sociopragmatics. Moreover, they roughly correspond with supercategories found in a number of grapholinguistic works (such as Venezky 1977; Rogers 1995). Notably, except for the linguistic fit, the mentioned fits can be applied to all three modules of writing systems.

Table 6. Criteria for the evaluation of writing systems and scripts

| Venezky (1977) | Coulmas (2009) | Cahill (2014) |
|---|---|---|
| — mechanically suited for the language it is to reflect (2)<br>— compatible with [...] its social-cultural setting (4)<br>— psychologically/pedagogically appropriate for its speakers (3) | — convenience (3)<br>— tools (3, 4)<br>— general applicability and linguistic fit (2)<br>— expressive power (1, 2)<br>— simplicity (1, 2, 3)<br>— stability through time (2)<br>— monochrome coding (2, 3) | — linguistically sound (2)<br>— acceptable to all stakeholders (4)<br>— usable (3, 4) |
| **Baroni (2011)** | **Daniels & Share (2018)** | **Bauernschmidt (1980)** |
| — maximum distinctiveness (1, 2, 3)<br>— size of the graph(em)ic inventory (1, 2, 3)<br>— cognitive salience (3)<br>— maximum naturalness (1, 2, 3, 4)<br>— inner consistency (1, 2) | — linguistic distance (2)<br>— spatial arrangement and non-linearity (1, 2, 3)<br>— visual uniformity and complexity (1, 3)<br>— historical change (2)<br>— spelling constancy despite morphophonemic alternation (2)<br>— omission of phonological elements (2)<br>— allography (1, 2)<br>— dual purpose letters (1, 2)<br>— ligaturing (1, 2)<br>— inventory size (1, 2, 3) | — linguistic factors (2)<br>— psycholinguistic factors (3)<br>  – „magic of written language“<br>  – native speaker reaction<br>  – optimal inventory of symbols<br>  – overuse of symbols<br>— sociolinguistic factors (4)<br>  – symbol value<br>  – adjustments for dialects<br>  – unity of language families<br>  – prestige, numbers, and so forth<br>  – established alphabets<br>  – government agencies<br>  – transfer value<br>— practical factors (4) |
| **Rogers (1995)** | **Smalley (1964)** | |
| — linguistic (2)<br>— psychological (3)<br>— cultural (4)<br>— technical (4) | — motivation for the learner (3, 4)<br>— representation of speech (2)<br>— ease of learning (3)<br>— transfer (2, 4)<br>— ease of reproduction (3, 4) | 1. systematic fit<br>2. linguistic fit<br>3. processing fit<br>4. sociocultural fit |

The first category, the 1) *systematic fit*, is purely descriptive. To evaluate the systematic fit of scripts, what is of concern are the visual features of basic shapes and their systematicity within a given script. The main question, here, is whether scripts, as visual systems, are coherent. This subsumes questions such as: How systematic are the relationships between a script's basic shapes? Are the visual features consistently spread throughout the units of a script or are there outliers? An example of such an outlier would be |J| in Roman script given that almost all other uppercase basic shapes are oriented rightwards. Furthermore, are there any systematic gaps, i.e. basic shapes that would theoretically be well-formed according to the visual features of a script but that are not actually part of the script? In this study, the focus will be on this graphetic systematic fit; however, the systematic fit can also be evaluated for graphematics and orthography. For them, the main question, too, is whether the relationships between their units – graphematic units (such as graphemes) in one case, orthographic rules in the other – are systematic. For example, Chapter 5 will introduce the simplification of Chinese characters as an example of a deteriorating orthographic systematic fit: Here, systematic relationships between the subcomponents of characters were opacified by an inconsistent simplification of characters. In a nutshell, orthographic intervention, while making individual characters easier to use, made the system less systematic.

Next, the 2) *linguistic fit*, which concerns exclusively the graphematic module, describes the nature of the relationship between a writing system and its underlying language.[198] In other words, it deals with the question of how well a writing system fits its language – in a strictly linguistic sense. This is done by evaluating the semiotic quality of the graphematic module as the link between the visual and the linguistic (cf. Baroni 2011). Universal (and rather general) preferences can certainly be determined for a "most natural" linguistic fit of any writing system; however, the linguistic fits of specific writing systems must be evaluated individually, as the interaction between a specific language system (especially its idiosyncratic features) and its writing system is subject to system-specific factors (cf. the levels of naturalness in the next section).

The linguistic fit is actually the subject of many grapholinguistic publications that describe how units of writing correspond with linguistic units. These relations are often termed *grapheme-phoneme correspondences* in the case of segmental phonographic systems. Given that these correspondences can be conceived of as semiotic relations, the semiotic naturalness parameters of Natural Morphology (cf. Section 4.2.3), one of the two core subbranches of Naturalness Theory, are expected to be fruitful in evaluating the linguistic fit. Indeed, several writing systems (or rather their graphematic modules) have already been evaluated with respect to two naturalness parameters, transparency and uniformity. Thus, a grapheme is *transparent* if its basic shape relates to only one linguistic unit,

198 Arguably, the linguistic fit is a special type of a so-called *semiotic fit* that evaluates whether a secondary semiotic system fits its primary semiotic system (in this case whether a writing system, specifically its graphematic module, fits the underlying language system). This semiotic fit can also be evaluated for the orthographic module. The central question, then, is whether orthographic rules fit the underlying graphematic regularities of the writing system or whether they are idiosyncratic and conceal the systematics of the graphematic module.

whether it is a phoneme, a syllable, a morpheme, or a different type of unit.[199] The German grapheme <v>, for example, is not transparent because the basic shape |v| is used for both /f/ as in <viel> 'much' and /v/ as in <vage> 'vague'. Inversely, a linguistic unit is *uniformly* represented if there is only one basic shape corresponding with it. This, for instance, does not hold for German /f/ which can be written as |f|, |v|, and a combination of |p| and |h|, <ph> (cf. Nerius 2007; for an elaborate analysis of the graphematic solution space of /f/, see Balestra, Appelt & Neef 2014; cf. also Section 2.3). If the relation between the basic shape and the linguistic unit is both transparent and uniform, the grapheme is *biunique* (cf. Munske 1994: 19f.). Complete biuniqueness can be evidenced, for example, by the notation system IPA (International Phonetic Alphabet): one basic shape correlates with exactly one sound and one sound is always written with the same basic shape.

The parameters of transparency and uniformity are commonly used to describe whether the graphematic/orthographic module[200] is *shallow* (as is the case in Finnish, for example) or *deep* (e.g. English) (cf. Katz & Frost 1992), representing the gradual distinction known as *orthographic depth* that can be utilized to help assess (part of) a given writing system's linguistic fit.

Share & Daniels (2016: 23–26; cf. also Daniels & Share 2018: 104–110) point out that this monodimensional concept of orthographic depth applies predominantly to European alphabets and therefore challenge its value for other types of writing systems. Indeed, the concept of orthographic depth only concentrates on the *phonological* biuniqueness of writing systems and thus focuses on phonographic writing systems. However, as a semiotic criterion, biuniqueness can also be evaluated for the relations between basic shapes and non-phonological linguistic units. Thus, for instance, the transparency and uniformity of graphemes in the Chinese writing system (which relate to morphemes) can of course also be evaluated. This reveals that Chinese graphemes are indeed largely transparent and uniform. By contrast, given that Chinese graphemes only offer (often opaque) clues to pronunciation, according to the phonology-centric orthographic depth hypothesis, the Chinese writing system is automatically discarded as "deep". Acknowledging this "monodimensionality", Daniels & Share (2018: 104–110) propose ten dimensions of orthographic depth: (1) *linguistic distance* (differences between spoken and written language), (2) *spatial arrangement and non-universality*, (3) *visual uniformity and complexity*, (4) *historical change* (retention of historical spellings

**199** Most writing systems are not purely of one type (e.g., alphabetic, morphographic, and so on) as different graphemes within the system can relate to different types of linguistic units.

**200** Mentioning both modules here is not an indication of a reluctance to commit but related to the question of what the attributes *shallow* and *deep* refer to. I argue that it can be both modules: the graphematic module is shallow if the graphematic relations between basic shapes and linguistic units are predominantly transparent, while an orthography is shallow if the prescriptive *standardization* of the writing system is transparent, i.e. its orthographic rules. Thus, in theory, a writing system whose graphematic module exhibits a high degree of transparency could still be deep because of idiosyncratic and unsystematic orthographic rules.

despite pronunciation change), (5) *spelling constancy despite morphophonemic alternation*, (6) *omission of phonological elements*, (7) *allography*, (8) *dual purpose letters*, (9) *ligaturing*, and (10) *inventory size*.

Not all of these dimensions can be neatly categorized as being relevant to the linguistic fit (cf. Table 6); some concern the graphematic processing fit and some are not concerned with the graphematic module at all but rather with graphetics (such as "visual uniformity and complexity"). Altogether, the dimensions proposed by Share & Daniels for a more universal and inclusive concept of orthographic depth roughly correspond with the considerations that I subsume under the label of grapholinguistic naturalness in the present approach. This is underlined by the fact that Share & Daniels (2016: 26) "regard these 10 dimensions as merely a catalyst for discussion of the multi-dimensional nature of writing system complexity". Notably, their dimensions are not systematically categorized and, like many of the criteria included in the other lists in Table 6, are predominantly inductive. Therefore, the authors admit that "[i]n several cases, [...] the dimension has yet to be addressed empirically" (Daniels & Share 2018: 104). I argue that a treatment of writing systems in a naturalist framework, which requires external evidence, is a systematic way of assessing a writing system's complexity, or, in naturalist terminology, its (un)naturalness.

While the systematic and linguistic fits are determined descriptively and without recourse to external evidence, this does not hold for the next two fits, whose relevance has been underlined, for example, by Venezky (1977, 2004). They shift the focus from the structure of writing systems to their use.

The 3) *processing fit* describes the relationship between writing systems and the human faculties necessary to process them: How suited is a writing system for the hands, the eyes, and the brain? What is conceptualized as *processing fit* here is, in large part, congruous with the traditional definition of naturalness in Naturalness Theory: it defines those structures as natural that are easier to process by humans (see next section). The following questions are central: Which features of the graphetic, graphematic, and orthographic modules make them easier or harder to process? Do the systematic and linguistic fits affect their respective processing fits? In other words, does a system's structure influence its users' performance in using it? If so, how? In theory, the processing fit could also be determined first and, based on the results, assumptions could be made about the systematic and linguistic fits. In any case, the broadest hypothesis concerning these fits' relationship is: the better the systematic and/or linguistic fits of a system, the better its processing fit. The same is expected to hold vice versa. When it comes to the question of cause and effect, the processing fit intuitively appears to be a consequence of the systematic or linguistic fits. However, the inverse relation should also be considered: in the process of the diachronic development of writing systems, users' (mainly unconscious, but partially also conscious) actions might lead to a change or even elimination of some features (in the vein of 'invisible hand' theories, cf. Keller 2014) and this occurs especially when features are not suited for processing needs. Accordingly, the processing fit acts as "human pressure" and can influence the systematic and linguistic fits (cf. also Dehaene 2009).

Not only purely descriptive as well as physiological, psychological, and neurobiological constraints contribute to (un)naturalness in writing but also important sociocommunicative, ideological, and cultural aspects. The 4) *sociocultural fit* deals with these very aspects. Cahill (2014) describes several non-linguistic factors that play seminal roles in the context of creating new writing systems for hitherto unwritten languages. They include the choice of either adopting an existing script or designing an entirely new script but also governmental policies that are enforced upon writing systems and, crucially, sociolinguistic factors such as attitudes towards varieties of one's own language or towards other languages and the associated wishes of signaling affiliation with or distance from a certain social group (cf. also Sebba 2009; Unseth 2005). Notably, whether a (new) writing system succeeds depends on the consideration of these factors, which is why Cahill (2014: 23) stresses "the importance of local community involvement" in the creation of new writing systems. If a community dislikes or even rejects a writing system that has been devised for its language, the system has effectively failed, even if its systematic, linguistic, and processing fits exhibit high degrees of naturalness. The fact that sociocultural factors often override other factors more than warrants their inclusion in studies like the present one – it makes it a necessity.

In conclusion, the criteria shown in Table 6 as well as the fits they are assigned to serve as important orientation tools that scholars can use to organize natural features of and in writing systems after they have been deduced from external evidence. In the following, starting with Chapter 5, entire chapters will be devoted to each fit.

## 4.2 Naturalness Theory

> [I]n general the conditions of the *use* of language [...] are responsible for the *nature* of language. (Stampe 1979: 43, emphasis in original)

Naturalness Theory is a functional linguistic paradigm that treats language as a tool for communication and cognition. Its story starts in the late 1960s when the notion of *naturalness* gained importance in linguistics with the advent of so-called *Natural Phonology* (in the following: NP), the first subbranch of what would later be subsumed under the heading of Naturalness Theory. NP was helmed by David Stampe, whose PhD thesis, originally titled *How I spent my summer vacation* (1972, University of Chicago) and later published as *A dissertation on Natural Phonology* (1979), can be considered the official birth of the theory. NP was then further developed by Stampe and other naturalists, among them Patricia J. Donegan, who shaped it with important contributions (such as Donegan 1978/1985). Some historiographic accounts additionally mention Charles-James N. Bailey as a co-founder of the approach and early naturalist; however, he focused on variation and language change instead of phonology (cf. Wurzel 1988: 99; Elsen 2014: 174). Even though there exists an extensive body of research within the paradigm of NP (cf. Luschützky's 1991 bibliography), overviews of the main positions of the theory are sparse. The few available accounts include Donegan & Stampe (1979) and Donegan & Stampe (2009), Dressler (1984), and, from the perspective of application in communication disorders, Edwards & Shriberg (1983). These works should be consulted for a more extensive descriptive picture of NP, as this subchapter merely aims to selectively highlight those historical circumstances and theoretical ideas that are expected to be relevant and fruitful for a functional theory of grapholinguistics.

Historically, and in terms of theoretical ideas, NP represents the "modern development of the oldest explanatory theory of phonology" (Donegan & Stampe 1979: 126). It is a direct response to the fact that "during the twentieth century, the rich fabric of explanation and evidence traditional phonology had woven of causality, intention, and consciousness, was dismissed as a tissue of unscientific reasoning" (Stampe 1985: 133). In essence, NP embodies a continuation of precisely these 'traditionally phonological' ideas that had emerged in the 19th and 20th centuries.[201] Incentives that were especially central for Stampe's motivation to establish NP were his observations that there exist cross-linguistic patterns in children's acquisition of phonology and that children's phonology is more complex than adults' (cf. Hurch 1988: 7). These considerations led Stampe to take up Sapir's (1933) phoneme theory, resulting in a mental(istic), psychologically defined phoneme concept that stands in sharp contrast to the structuralist and generativist conceptions of the phoneme. Here, it is noteworthy that the phoneme is not a particularly central concept in NP – this role is served by phonological processes instead (see below). While structuralists reduced the phoneme to its distinctive function – its opposition with other phonemes – generativists, particularly Chomsky and Halle, did not accept this 'discovery procedure' as a definition of the phoneme, and, because of Halle's phonological analysis of Russian, rejected the phoneme altogether[202] (cf. Stampe 1985: 133; Nathan 2007: 93). Naturalists,

201 Specifically, Donegan & Stampe (1979: 126) state that the elements included in NP "evolved in nineteenth-century studies of phonetics and phonetic change (Sweet, Sievers), dialect variation (Winteler), child speech (Passy, Jespersen), and synchronic alternation (Kruszewski, Baudoin), and developed further, still without integration, in twentieth-century studies on dynamic phonetics (Grammont, Fouché) and phonological perception (Sapir, Jakobson)".

202 This is based on Halle's (1959) finding that if three levels of representation are assumed, a (1) morphophonemic level, a (2) phonemic level, and a (3) phonetic level, the process of regressive voicing assimilation in Russian must occur separately (and thus, twice) on two of those levels: an assimilation of /k/ to /g/ is morphophonemic, because both /k/ and /g/ are phonemes of Russian – thus, one phoneme changes into another. However, phonemes whose voiced/voiceless counterparts are *not* phonemes of Russian also take part in this assimilation. The phoneme /ʧ/, for instance, becomes [ʤ], which is itself not a phoneme but an allophone. Thus, this latter assimilation is not a morphophonemic but a phonemic process. Given that the same process would

too, considered the 'objective' structuralist definition to be reductionist[203] but did not opt out of the phoneme altogether. Instead, in NP, phonemes are conceived of as mental sound intentions shared by speakers and hearers. In a departure from rules and formal levels of phonological representation and including a categorical distinction between 'pure' phonology and morphonology, NP's core tenet is that there exist phonological processes that are applied to eliminate phonetic – articulatory as well as perceptual – difficulties for speakers and hearers. Since they are determined by human physiology, or human 'nature', they are termed *natural* phonological processes and lend the approach its name.

After the establishment of NP, *Natural Morphology* (in the following: NM) was developed as Naturalness Theory's second major subbranch. Eventually, it would be perceived by many as its "most significant achievement" and "the one [subbranch, D.M.] which has been best worked out" (Gaeta 2006: 8). It was founded by Austrian linguist Wolfgang U. Dressler and German linguists Willi Mayerthaler and Wolfgang U. Wurzel.[204] Accounts of when exactly NM was first established differ, but Dressler himself dates the inception of the theory to 1977 (cf. Dressler & Kilani-Schoch 2016: 356; Dressler 2006: 539) and states that it is based on a chapter in Mayerthaler (1977). The official and public 'birth' of the theory is said to have taken place at the 1979 *LSA Summer Institute* in Salzburg, Austria (cf. Kilani-Schoch 2001: 234).

While the development of NM was given crucial impetus by NP, which is underlined by the name adopted for the then-nascent theory,[205] and it is explicitly considered "a semiotically and cognitively based functionalist theory in the continuation of Natural Phonology" (Dressler & Kilani-Schoch 2016: 356), the distinct characteristics of morphology and phonology required integrating additional ideas. The first notable difference between NM and NP concerns their theoretical roots: whereas NP criticized structuralism in a number of respects, NM relies heavily on its "structural heritage" (cf. Wurzel 1988). Sapir is named as an important influence in both NP and NM, and additional linguists who are mentioned for NM include the neo-grammarian Hermann Paul as well as August Schleicher and Vladimír Skalička. However, the arguably most important groundbreaker for

have to apply twice on different levels, Halle rejected the phonemic level and the notion of phoneme (cf. Dresher 2011: 257f.; Schane 1971: 517–519).

203 Consider Donegan & Stampe's (1979: 129) reckoning with the distinctiveness criterion: "But words are not only distinguished by sounds, they are made up of them. It is no less important that the sounds that constitute words be distinguishable than that they be pronounceable, combinable, and perceivable (articulate, audible)".

204 At times, Austrian linguist Oswald Panagl is additionally mentioned as a fourth founder (cf. Kilani-Schoch 2001: 234). Together with Dressler, Mayerthaler, and Wurzel, he authored the programmatic *Leitmotifs of Natural Morphology* (1987).

205 Dressler was an associate professor at the Ohio State University in 1970-71; he returned there as a guest professor in 1977 (cf. https://www.oeaw.ac.at/fileadmin/Institute/ACDH/img/Team/CV_WUD_deutsch.pdf, June 4th, 2020), which was the time when David Stampe and his colleagues were actively working on NP there (Bernhard Hurch p. c.).

NM was Roman Jakobson (cf. Wurzel 1988: 103), and more generally, the Prague School of linguistics. Some even regard NM – at least in part – as a product of said Prague School (cf. Lieber 2014).

Perhaps Jakobson's greatest achievement of relevance for NM is his treatment of semiotics that eventually led to the formulation of a semiotic metatheory for Naturalness Theory (cf. Dressler 1999; Crocco Galèas 1998: 8–10). Accordingly, some scholars observe that "[s]tructuralism has been united with semiotic studies by Natural Morphologists" (Beard 1994: 2576). The implementation of a semiotic metatheory that relies predominantly on Peircean semiotics is based on the assumption that the nature of semiotic relations bears on the cognitive processing of signs and sign systems, the central of which is of course language (cf. Dressler 1987: 165). In short, one of NM's central claims is that several features of the semiotic structure of a sign – understood as the relationship between its signans and its signatum – bear on how humans cognitively process said sign. Based on this assumption, several naturalness parameters (such as transparency and uniformity) were described; these are grounded in the features of semiotic structure and are said to be cognitively 'real'. An additional aspect of Jakobson's work that proved relevant for NM is the concept of *markedness* that also originated in the Prague School. In NM, the term *naturalness* is interpreted as a conceptual and terminological synonym of *unmarkedness* (cf. Dressler 2000: 288) – as such, it is considered the "diametrical opposite of markedness" (Wurzel 1994: 2591).

Its integration of semiotics on the one hand and markedness on the other hand means NM shares similarities and even partially overlaps with other theories or linguistic schools of thought, including *Cognitive Linguistics* (cf. Dressler 1990, but also Nathan 1999 for a discussion of the compatibility of Naturalness Theory and Cognitive Linguistics), modern *Markedness Theory*, and *usage-based approaches* to linguistics, a paradigm that was initially developed by, among others, Joan Bybee [Hooper], who had also treated questions of morphological naturalness.

Concerning an operationalizable definition of linguistic naturalness in NM, Dressler (2000: 288) specifies that *natural* is "often synonymous to cognitively simple or easily accessible", while Mayerthaler concludes that the meaning of "more or less natural [...] really boils down to '*more or less easy for the human brain*'" (Mayerthaler 1987: 27, emphasis in original). In these quotes, the focus is on cognition; however, physiology – e.g. what is easier or less easy for the articulators and receptors, mouths/hands and ears/eyes – and social factors – e.g. what is most natural for the purposes of communication – are equally important. From this follows that what is *more natural/less natural* cannot be evaluated (exclusively) language-internally but requires the consideration of language-external evidence. Furthermore, it cannot be evaluated in isolation but only in comparison: nothing inherent in a single linguistic element can reveal whether it is more or less natural than some other given element. Naturalness Theory relies heavily on the study of linguistic *performance* (cf. Dziubalska-Kołaczyk 2002: 104), since the cognitive/physiological/social naturalness of linguistic phenomena, in other words, the ease with which they are produced and perceived, can only be determined when their *use* is considered.

In a nutshell, a comprehensive interim definition of linguistic naturalness reads as follows: *naturalness* refers to the effort involved in using language – this includes production as well as perception – with respect to external constraints. On the one hand, these external constraints are the physical and cognitive makeup of the human body, specifically the parts that are relevant for the use of language. On the other hand, psychosocial considerations of humans as socio-communicative beings are also crucial for linguistic naturalness. Accordingly, Bailey (1984: 229) speaks of "a balance between (bioneurolinguistic) structures and (sociopragmatic) communicational functions". This corresponds perfectly with Dressler's (1980: 75, my emphasis) claim that "[n]aturalness must be derived from considerations of the *nature of man*, who is not only a speaker-listener, but also a non-verbally communicating being conditioned by biological, psychological and social properties. Therefore[,] any 'natural linguistics' must be based on such extralinguistic considerations [...]". Crucially, Wurzel (1994: 2592) adds to the biological and social factors the aspect of culture-specificity that will play a pivotal role in a functional analysis of writing.

Naturalness Theory consists of three levels: in addition to the universal level (the only level that is treated explicitly in NP) that gives rise to a system-independent preference theory, NM features two more levels and corresponding subtheories: the typological and system-dependent levels and subtheories of naturalness. The former evaluates naturalness in types of language (such as agglutinative, isolating, ...), the latter what is natural in individual systems, which essentially boils down to the prioritization of the different naturalness parameters (cf. the next section).

As mentioned above, aside from cognition, socio-communicative factors prove crucial in NM: since language is a tool for communication, linguistic behavior is seen as a means for achieving social goals. This adds another layer of analysis, as the semiotic structure of signs not only has to be evaluated denotatively but also connotatively: linguistic structures always reveal additional information, e.g. about the speakers/writers producing them or the situation in which they were produced, which is obviously information that goes beyond mere propositions. Under this perspective, even phonology, despite dealing with 'meaningless', i.e. non-semiotic phonemes, can become a semiotic affair: speaking sloppily, i.e. not pronouncing an utterance carefully, is possibly not motivated solely phonetically, i.e. physiologically (for example due to fatigue or intoxication), as there can also be a semiotic motivation: If a person speaks sloppily, this might also be due to the fact that the speech situation is informal and he or she is talking to someone who is very familiar. In such a situation, the phonetic output can be semiotically charged, becoming a sign of the speech situation, the relationship between the interlocutors, etc.

External evidence is central to Naturalness Theory, where it counts as 'substantial' evidence: data from e.g. language acquisition, language disorders, and language change are considered to evaluate both natural phonological processes in NP and naturalness parameters in NM. The different underlying motivations of naturalness in the two subbranches – phonetics in NP, semiotics/cognition in NM – reflect directly the differences between phonology and morphology and result in varying understandings of 'natural'. While NP does not define naturalness as explicitly as NM, it is implied that naturalness is treated as an absolute attribute: everything that is phonetically realized by humans is phonologically natural since for it to have been materialized, it has had to run through several phonological processes that eliminated 'unnatural' obstacles. For NM, on the other hand, naturalness is a scalar, gradual concept: on each parameter, different degrees of naturalness can be evaluated ranging from more natural to less natural, for example biuniqueness—uniqueness—ambiguity. A phenomenon x is thus always evaluated in relation to another phenomenon y with respect to parameter z. Even though it is only explicated in NM, NP actually also works with a gradual rather than an absolute reading of naturalness: in the absolute reading, *everything* that occurs in language is seen as 'natural' since humans can process it. In a gradual reading, everything occurring in language can be compared and evaluated as more or less natural. Natural, in this gradual sense, means 'easier to process for humans', and this includes physiology (in NP), cognition (in NM), and social factors (in both). For the most part, naturalness is evaluated locally, i.e. for small-scale linguistic phenomena (such as words, phonological clusters, etc.) and with respect to given parameters, and not globally, with respect to whole language systems and the sum of parameters. However, as findings in linguistic complexity research suggest, the naturalness/complexity of whole systems *can*, in theory, be evaluated (as so-called *global complexity*, cf. Miestamo 2008; Szmrecsanyi & Kortmann 2012: 8–10), even if this represents a challenging endeavor.

Figure 26 gives a schematic, non-exhaustive overview of the theory: it illustrates the semiotic metatheory as well as other influential theories and paradigms (preference theories, functionalism in general), the extralinguistic bases of naturalness (cognitive, physiological, psychological, sociopragmatic), the three subtheories (universal, typological, language-specific), the linguistic subdomains to which the theory can be applied (phonology, morphology, syntax, text, etc.), and finally, the external evidence that is crucial for the theory.

In the following, I will deal with the question of how certain above-mentioned cornerstones of Naturalness Theory can be operationalized for and transferred to a functional theory of writing. Some of them originate from NP, others from NM. Notably, most of these cornerstones can be viewed from the different perspectives that the three modules of writing systems – graphetics, graphematics, and orthography – give rise to.

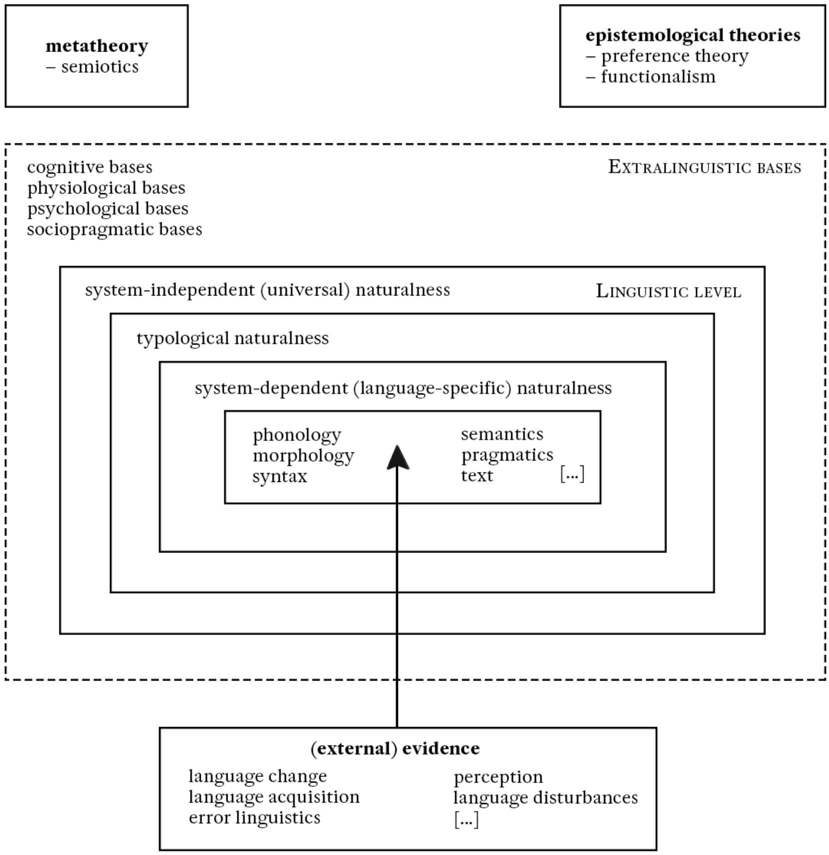



FIGURE 26. Structure of Naturalness Theory, translated and adapted from Miret (2009: 182)

### 4.2.1 Levels of naturalness

NM consists of three subtheories: *system-independent naturalness, typological naturalness*, and *system-dependent naturalness*. They can be transferred productively to a functional theory of grapholinguistics. Notably, questions concerning all the fits – systematic, linguistic, processing, and sociocultural – can be asked at all three levels, although the fits appear to correlate predominantly with one level, respectively (see below).

The level of *system-independent naturalness*, also referred to as universal naturalness, investigates the question of what is universally preferred. For the graphetic module, possible questions include: What are the universally preferred visual configurations for perception as well as motoric programs for the production of basic shapes (cf. Watt 2015)? What are the universally preferred features of a script that render it a coherent visual system? How do literate communities decide that scripts are more or less suitable for their specific sociocultural environments? Evidently, some of these questions are more reasonably, though not exclusively asked at a universal level (questions of the systematic and processing fits) than others (problems pertaining to the sociocultural fit). The latter might be better located at a lower level, i.e. the typological or even the system-specific level at which the fit of a given script for a given culture can be determined in a fine-grained manner by considering highly idiosyncratic culture-specific factors that arise in individual contexts.

Regarding system-independent naturalness in the graphematic module, major questions are: What is the preferred semiotic structure of graphemes? What is the preferred assemblage of graphemes in a grapheme inventory? These two questions could be subsumed under the very broad question: What is the most natural relationship between the visual and the linguistic? When the focus shifts to processing of written structures, the major question is how the semiotic structures influence cognition. Material preferences regarding basic shapes have already been assessed in the investigation of naturalness in the graphetic module, and the naturalness of the specific linguistic units that graphemes correspond with – specific phonemes, syllables, morphemes, etc. – is not addressed by a functional theory of writing but by the other components of Naturalness Theory (NP, NM, and others such as Natural Syntax). However, the question of whether one *type* of linguistic unit that graphemes potentially correspond with is universally preferred is a question of interest, and it is noteworthy that a "primacy of the syllable" (Daniels 2018: 12) has been postulated, claiming that syllables are more natural as processing units than segmental phonemes. From the perspective of processing, this alleged primacy will be addressed in Section 7.2.1.

The second subtheory, *typological naturalness*, deals, at the graphematic level, with the relationship between types of language (isolating, agglutinating, etc.) and types of writing systems (phonography and its subtypes vs. morphography). Here, Halliday's ([1977] 2010: 103, my emphasis) oft-cited quote "[i]n the course of this long evolution, a language usually got the *sort* of writing system it deserved" comes to the forefront. What he most likely refers to with *sort* is a writing system's *type*. Notably, compared with the number of language types that have been described, types of writing systems are strikingly sparse. The basic dichotomy between phonography and morphography becomes only minimally more nuanced with the establishment of several subtypes of phonographic systems (cf. Section 2.7), but then again, as Weingarten (2011) noted, writing system typology might still be in its infancy – much like many subfields of the underdeveloped grapholinguistics. So, is Halliday correct in his opinion? Do languages get the type of writing system they "deserve"? Do features of language types preferentially correlate with certain features of types of writing systems?

Graphematic questions at the typological level concern predominantly the linguistic fit but they do also venture into the processing and sociocultural fits. One prominent example is the question of whether the morphographic Chinese writing system could be written with an alphabet (cf. Meletis 2018: 72; Rogers 1995: 39; DeFrancis 1943; cf. Section 6.1). Many argue that because of the isolating nature of Chinese and the sheer abundance of homophony, writing Chinese alphabetically would lead to numerous homographs and thus ambiguities that would render the writing system less natural from a processing perspective. Absurd situations would arise in which an identical graphematic word could be reproduced several times in a minimal context, with each instance representing a different morpheme, i.e. a different meaning.[206] Precisely because Chinese is an isolating language, i.e. there is no inflectional information that needs to be graphematically encoded, a morphographic writing system in which graphemes correspond directly with morphemes is clearly the most economical solution. Finally, the question of why the Chinese have not opted to replace their system with an alphabet is at its core also a deeply cultural question (cf. what DeFrancis 1943 calls "the social problem"). To give a simplified answer, adopting a different script (and with it a distinct type of writing system, a phonographic one) would mean drastically cutting ties with thousands of years of cultural tradition, a decision that could be interpreted as bowing to the pressures of the West and "admitting" that the West and its inventions – such as the alphabet – were superior. This superiority, of course, is highly questionable, which might be an additional reason why the Chinese hold onto their traditional script and their morphographic writing system. Also, one of the well-known main advantages of morphography is that it enables the mutual intelligibility of different Chinese varieties and politically unites them even when spoken pronunciations are not mutually intelligible (cf. Chen 2004: 114–128). If a strictly or even predominantly phonographic writing system were to be employed, this advantage would be eradicated.

Since no typology of scripts exists yet (but cf. Section 1.3), it is difficult to imagine how typological graphetic naturalness could be assessed. If a base criterion for a typology were to be identified, for example, *roundness* vs. *angularity* of a script's basic shapes, questions pertaining to the resulting types could be examined, such as: Is the round type of scripts (as evidenced by, e.g., Georgian script, Telugu script) more easily read and written than the angular type of scripts (e.g. Chinese, Korean, etc.)? Since no such types have been assumed and described yet and this present book can merely uncover possible reasonable choices for typological base criteria, investigations of questions such as these must be postponed to a later stage of a functionalist theoretical grapholinguistic enterprise.

The last subtheory described by NM is *system-dependent naturalness*. At this level, the main central question is what is natural in a given script or writing

206 Chao (1968: 120f.) provides the extreme example of a story consisting of 36 instantiations of the syllable *xi* with one of the four tones. He notes that "[i]t makes absolutely no sense when read aloud in modern Mandarin, but from the writing a reader of classical Chinese can make out the story [...]". Rogers (2005: 29f.) cites another example of a short story consisting only of morphemes/words that have as their phonological representation the syllable *shi* with one of the four ones.

system. This question, too, can be assessed with respect to all of the fits. Graphetically, questions pertaining to individual scripts are: how natural are the features of a given script? Are they in conflict with each other? From a sociocultural perspective, what are, in a specific context, cultural, ecological, technological, social, and political factors (cf. Downey 2014) that influence the choice of a script for a given writing system?

Graphematically, each writing system offers a unique set of system-dependent naturalness values on the parameters that were identified as crucial at the system-independent level. At the system-dependent level, which overrides the typological level, an interesting observation is that there exist no "pure" writing systems. Type-mixing is almost always involved, rendering questions of how individual systems choose to solve specific problems crucial: Why, for instance, does German opt to give the morphographic principle some weight, writing <Kälte> 'the cold' instead of <Kelte> 'Celt' when graphematically representing the noun that derives from the adjective <kalt> 'cold'? And why is this morphographic principle not as important in Finnish, which is likewise an alphabet? At the system-specific level, the linguistic fit of individual writing systems can be studied, assessing how well a given writing system fits the language it is used for. Notably, the processing fit of a given system might also deviate substantially from what is universally or typologically natural in processing: take Thai, which is an abugida (or, per Gnanadesikan 2017, a mostly vowelled āksharik/linear segmentary, cf. Section 2.7) but deviates from many other more prototypical abugidas. Specifically, in Thai, not all vowel graphemes are smaller in (relative) size than consonant graphemes, and those that are of equal size occupy their own segmental spaces rather than just attaching to consonant graphemes within their respective segmental spaces. Additionally, some vowel graphemes spatially precede consonant graphemes although, in the phonological structure that the graphematic structure corresponds with, the vowel phonemes temporally follow the consonants – a phenomenon known as *misaligned vowels* (cf. Winskel 2009). These idiosyncratic features take a toll both on the linguistic and processing fits of the Thai writing system, which has a dramatic effect on its system-dependent naturalness. Last but definitely not least, with respect to system-dependent graphematic naturalness, the sociocultural fit takes center stage: which are the context-specific sociocultural factors that significantly affect how a writing system for a yet unwritten language is designed?

These three levels of naturalness interact with each other. Crucially, lower levels have the potential to override higher levels: thus, typological naturalness can override system-independent naturalness and system-dependent naturalness can override both system-independent naturalness and typological naturalness. While universal preferences are paramount mostly in the search for general explanations of why certain structures and phenomena recur across writing systems, lower levels are needed to explain more specific structures and exceptions. In other words, the system-independent level and, to some degree, the typological level are concerned with the unity of writing systems, i.e. their shared core, while the system-dependent level captures the diversity of individual scripts and writing systems. An additional level was addressed but not given an elaborate treatment in NM: norms. The increase of the power to override continues through this level, as norms can override all higher levels. One reason why this level might have remained undeveloped in the context of NM is that it is not straightforwardly clear how naturalness should be defined for it.

This situation can change with a naturalist treatment of writing, where norms are central: the existence of the (optional) orthographic module of writing systems highlights the importance of the level of norms for a grapholinguistic theory. Norms effectively superimpose both the graphetic and the graphematic modules. In graphetics, norms exist mainly in the form of conventions that remain mostly implicit. For instance, many people feel that it is inappropriate to use a ‘childish’ typeface such as Comic Sans when designing a resume (cf. Meletis 2020a). However, norms might also be explicit in the form of orthographic rules: in Chinese, forgetting or misplacing a stroke, producing a wrong stroke, omitting a stroke in a character, or producing the strokes in the wrong sequence are all orthographic mistakes at the graphetic level (cf. Law et al. 1998). They might result in the production of an existing basic shape that is part of a different graphematic relation – in this case, a graphetic mistake simultaneously constitutes a graphematic mistake. In Arabic, too, when a dot is omitted or misplaced, this can result in a different basic shape than the one that was intended, for example when |د| is written instead of |ذ|, two basic shapes that take part in different graphematic relations (cf. Brosh 2015). Thus, although it lacks visual salience, the dot can serve as a distinctive visual feature in Arabic script – in examples like this even the only one. Consequently, if it is omitted, a different grapheme will be produced and the word is misspelled. Such minimal graphetic differences between two distinct shapes are, thus, a source of mistakes.

More so than the graphetic module, norms superimpose the graphematic module, specifically its linguistic fit: explicitly codified orthographic norms such as entries in dictionaries restrict the graphematic level by selecting one or (more seldom) more variants from inside the graphematic solution space as officially correct. In some cases, that variant is not even part of the graphematic solution space, i.e. it is a spelling that is deemed correct although it is not even licensed by the graphematic module (cf. Chapter 3). This is possible since norms – at least with respect to the standardization of the graphematic module – are generally determined externally in that they are not imposed upon a system by the users of a writing system in the vein of an invisible hand (cf. Keller 2014) but instead by authorities of linguistic policy such as language academies, ministries of culture, or other types of institutions. However, prototypically, orthographic rules are, to a large degree, based on the conventions upon which language users have initially and implicitly agreed, even if they are ultimately decided on by authorized commissions (cf. Neef 2015: 716; Dürscheid 2000). In the case of German orthography, for instance, the actual use of the system is observed and taken into account by the Council for German Orthography (cf. Güthert 2016: 16–20).

Even though facilitating literacy acquisition is a main goal of many orthography reforms, a good processing fit is not always the primary motivation guiding the design of orthographic rules, which is why several of them appear arbitrary and idiosyncratic. In fact, orthography often acts as a kind of wall as it greatly complicates the search for natural features in the graphetic and graphematic modules behind it. This is due to the fact that the orthographic surface that users produce does not necessarily reveal what is going on in their minds as they might only, given orthography's social bindingness (cf. Section 3.1), obey the rules. Therefore, violations of orthographic norms, whether they are conscious or unconscious, are central to the investigation of grapholinguistic naturalness (cf. for methodology and evidence also below).

### 4.2.2 Natural processes

In the transfer of NP's ideas to a functional theory of writing, the most imminent question is how the concept of *natural processes* can be reinterpreted in order to be applied to the written modality of language. So, in this section, I will investigate precisely the question of whether there exist natural processes in writing and if so, how they compare to natural processes in phonology. Before I proceed, I want to note that – as the term natural phonological *process* underlines – NP focuses on the dynamic processes of speech production and only secondarily on the output itself. The same applies to processes involved in written production. They are always intricately linked to human physiology and cannot be described independently of human production and perception. This means that they are always a matter of the processing fit.

An example of a natural phonological process is the devoicing of final obstruents, as in German *Hund* 'dog', the phonological representation of which is (in NP) /hʊnt/ – here, the feature [+voiced] is substituted by the feature [−voiced] in the final obstruent because voiced obstruents are 'more difficult' to produce (cf. Hurch & Nathan 1996: 235). NP's main claim is that natural phonological processes, as sound substitutions, are reactions to difficulties in the articulation and perception of speech. In L1 acquisition, children must inhibit certain processes that are not active in the phonological system of their language. A child acquiring English, for instance, must inhibit the natural phonological process of final obstruent devoicing, because, in English, final voiced obstruents do occur. The phonology of a given language is thus shaped by which processes are inhibited and which remain active.

In NP, natural phonological processes are contrasted with *morphonological rules* which are not considered to be natural in the same way. An example of such a rule is umlauting as in German SG *Mann* /man/ 'man', PL *Männer* /ˈmɛnɐ/ 'men'. The most important difference between natural phonological processes and morphonological rules is that processes are phonetically motivated while rules represent morphophonemic alternations (cf. Hurch 2006: 541). Because they are conditioned phonetically, processes are, as mentioned above, automatic in that they almost always apply as long as a specific phonetic context is met – such as when obstruents occur in syllable-final position. Rules, by contrast, are not automatic; their application is not motivated phonetically but lexically or grammatically (cf. Hurch 2006: 541; Hurch 1988: 8–9). Processes, thus, have a phonetic motivation, whereas rules do not need to have such a motivation but certainly *can* have it (cf. Donegan & Stampe 1979: 144), the latter often being the case when processes have developed into rules diachronically. In these cases, phonetic motivations might still be transparent in rules. Furthermore, processes can be seen as functioning bottom-up, as speakers' limitations shape the phonological system of a language. They are "restrictions the speaker imposes on his language", while for rules, the opposite is true: They are top-down restrictions "the language imposes on the speakers" (Hurch 2006: 541). Processes are productive, rules are not (cf. Hurch 1988: 9; Dressler 1984: 38). As expected, thus, processes apply in speech errors and tongue-slips, in foreign words, etc. (cf. Donegan & Stampe 1979: 144). They also lead to negative language transfer, when a process of an L1 is applied in the context of an L2 in which it is inhibited (or vice versa). For example, speakers of German must suppress final obstruent devoicing when learning English; failing to do so leads to mistakes in pronunciation. Rules, on the other hand, are not transferred from L1 to L2 (cf. Donegan & Stampe 2009: 12). A possible reason for the negative transfer of processes in L2 acquisition is that violating a process "requires special motivation" (Donegan & Stampe 2009: 5) as well as "special attention and effort" on behalf of the speaker. Voicing final obstruents in an L2 when in the speaker's L1, they are devoiced, does not come 'naturally', it requires effort. The violation of rules, by comparison, does not require special effort (cf. Hurch 1988: 10).

The central dichotomy of types of natural processes in NP is based on their causality: fortitions (also strengthening or foregrounding processes) are distinguished from lenitions (also weakening or backgrounding processes). While fortitions are centrifugal and enlarge the perceptual distance between sounds – meaning more effort is required for the speaker to produce a sound – lenitions are centripetal and "embrace all assimilatory tendencies which make pronunciation less expensive" (Hurch 2006: 542), rendering an utterance more difficult for a listener to understand. These two types of processes can be traced to the antagonism between keeping to a minimum the efforts for the speaker in production vs. for the hearer in perception. A parallel situation is found when considering the conflicting needs of writers vs. readers.

To return to writing, is there an analog to processes in the production of writing, i.e. are there natural graphetic (and possibly graphematic, see below) processes that arise from the difficulties of the physical act of writing? Indeed, evidence suggests that some sequences of basic shapes are harder for children to produce in handwriting than others (cf. Gosse et al. 2018; cf. Section 7.1.2). In general, handwriting movements (as studied by *graphonomics*, cf. Kao et al. 1986) can be viewed through the lens of naturalness, with the central question being which basic shapes or sequences of basic shapes[207] require less effort in production/are easier to produce. In turn, these differences in the effort involved in production might or might not act upon the basic shapes of a script and the *graphetic* (and in turn, *graphematic*) *graphotactics* of a given writing system. In theory, given basic shapes or basic shape sequences that are (mostly unconsciously) dispreferred by users might be avoided, and this could, consequently, lead to changes in the system. The described difficulties in the production process of certain graphetic sequences are more or less an exact analog of the difficulties that arise from producing sequences of sounds in speech. Here, one feature of a phoneme in the sequence is changed to render an utterance more easily pronounceable, such as in German *Hund* /hʊnt/ 'dog', where, as mentioned above, the feature [+voiced] is substituted by the feature [−voiced]. In writing, however, external tools are necessary, so a variety of tools and also different modes of writing and their interaction with the human hands need to be considered. Thus, typing on a keyboard of course also counts as writing even if it is, physiologically speaking, a process that differs fundamentally from writing with a pen. How can naturalness be evaluated in typing? A question that appears trivial at first glance but must definitely be taken into account in a naturalist evaluation of typing is the placement of individual keys on a keyboard: With respect to conventional English keyboards, is the QWERTY layout natural (cf. Noyes 1983; cf. also Section 7.1.2)?

Let us consider whether the features of natural phonological processes apply also to natural graphetic processes. Arguably, processes such as the ones involved in connecting basic shapes in cursive script are, just like natural phonological processes, *unconscious*. In general, natural graphetic processes are *physiologically conditioned* (subsuming as relevant factors human biology, writing instruments, and the writing surface). With respect to human endowment required for writing, natural graphetic processes are *innate*. The difference in ontology – that speech is "inborn" while writing must be learned through instruction – does not change the fact that pre-existing biological conditions are innate in both cases. The difference is that writing – unlike speaking – is, as mentioned above, not dependent solely on the innate physiological conditions of the human hands and eyes but also on the tools that are used for writing. Consequently, if certain natural graphetic processes are caused only by limitations of the hand, they are indeed *innate*. If, however, processes are caused by external factors such as the pen or the keyboard that is

207 A question that is relevant in NP is pertinent for graphetic naturalness, too: Is it possible to evaluate isolated units, i.e. units without a context? I believe it is, which will be shown in the discussion of the graphetic processing fit (cf. Section 7.1). Basic shapes are not restricted in their materiality the way as phonemes, which are constrained by the possibilities of the human physiology involved in speech production. Thus, the number of possible basic shapes is probably infinite while the number of phonemes that can be produced by humans is most certainly finite. The different features of individual basic shapes – e.g. their degree of graphomotoric complexity – strain human capacities to different degrees. In other words, natural graphetic processes can occur within individual basic shapes, meaning that these can be compared with respect to which and how many natural graphetic processes they evoke.

used and could, therefore, be avoided if different tools were used, they are not innate. For instance, if typing on a keyboard with a specific layout results in a more natural writing process than typing on a keyboard with a different layout, then this is a consequence of the nature of the tool(s). Tools, of course, always interact with the innate physiological limitations of the human hands, but the crucial point is that ensuing natural graphetic processes are not caused solely by innate human biology. Commonly, whenever a given graphetic context is met, natural graphetic processes are applied, which means they are *automatic*. However, just like natural phonological processes, natural graphetic processes, while they are universal, are not applied universally: They are *variable* and thus, in a way, *optional*. An example: When writers are tired (cf. Parush et al. 1998), drunk, or write particularly fast (cf. van Drempt, McCluskey & Lannin 2011), to name only a few factors, this potentially affects their writing and the natural graphetic processes they apply or, crucially, fail to apply. An important consequence of the fact that the application of processes is not obligatory is that the application or non-application has the potential to be sociosemiotically charged. Thus, even if natural graphetic processes are primarily conditioned physiologically, they can also be caused by sociopragmatic reasons, for example, when a scribe – out of respect for the addressee, for example (see below) – attempts to write in an especially legible (or even aesthetically pleasing) manner. In that case, the material appearance of writing becomes a semiotic affair.

While natural graphetic processes are automatic and conditioned physiologically, in writing, there also exist analogs to the morphophonological rules described in NP. In alphabets that exhibit a case distinction between lowercase and uppercase basic shapes, the initial grapheme in a graphematic sentence is materialized by an uppercase basic shape. This is a graphematic rule. It is not conditioned physiologically, and one could very well imagine basic shapes at the beginning of sentences being lowercase (as in the Georgian alphabet which lacks a case distinction or in informal writing in which capitalization is neglected). As discussed in the context of allography in Section 2.3, the sentence-initial majuscule is conditioned by linguistic factors external to the writing system. Thus, capitalization is not a natural graphetic or graphematic process. It is a rule.

The natural phonological processes described in NP were of articulatory nature. However, as Dressler (1984: 33) argued, perception is primary in speech while production is only secondary. A similar claim has been made for the written modality (cf. Primus 2006: 10), partially because members of literate societies are more often readers than writers. Furthermore, production almost always includes a feedback process of reading what is being written or has been written. This primacy of perception raises the important question of whether there are natural graphetic processes that are perceptual. An example could be the successful identification of an unrecognizable graph in a written word with the help of the available context.[208] An aspect that distinguishes productional natural processes

208 Strictly speaking, this would be a natural graphematic process, since it is not the graphetic context – i.e. only the visual appearance of the surrounding graphs – but the

as described in NP sharply from these proposed perceptual processes is that productional processes change the output. Perceptual processes, on the other hand, do not change the product that is being materialized but always apply on an *existing* or *emerging* output. However, I argue that a process is not defined by how it affects linguistic output, but precisely by how it affects articulation and perception. In this view, perceptual natural processes *do* also change the product, but a different kind of product: They do not alter the actual materialized product – whether acoustic or visual – but the mental product in the mind of the hearer or reader. This corresponds with Donegan & Stampe's (2009: 26) claim that hearers do not hear what speakers actually produce but what speakers intend to say. Thus, if a speaker were to "fail" to apply an articulatory natural process, as they are variable (and the speaker might be tired or drunk, see above), said process might still be applied in perception by the hearer.[209] I argue that the same holds for writing: When someone produces an utterance rather illegibly and the reader has to make more of an effort in deciphering it, the writer's facilitation-based lenitions require fortitions on the reader's behalf.

Given the richness of visual variety in writing, in the grapholinguistic conception of natural processes, the horizon must be widened: For example, printing a portion of a text bold, underlining it, or setting it in a larger font size during the digital production process could be interpreted as (conscious) graphetic fortitions that are simultaneously graphematic fortitions (see below). Graphetic fortitions can generally be defined as processes – whether conscious or unconscious – that make a written utterance more *legible* for the reader. Note the crucial difference between *legibility* and *readability* here: text needs to be legible on a material (i.e. graphetic) level to be visually *recognized*, whereas it needs to be readable on a conceptual (i.e. graphematic) level to be *understood* (cf. König 2004: 18; Lund 1999: 15–20). Legibility is an important part of readability. In short: When writers attempt to make their writing more readable (see below), the resulting process is not a graphetic, but a graphematic fortition.

Graphemes are signs, phonemes are not. As conceptualized in NP, phonemes are substantial, i.e. material linguistic units. By contrast, in the pres-

*graphematic* context and the linguistic information that it provides that allow identifying the graphematic status of the unidentified graph, i.e. the graph is identified top-down. If a certain graph in a given word is completely unrecognizable, it might never be classified as a member of a given basic shape (as the visual features required for that are just not there), but even then, because of the graphematic context, the abstract *grapheme* it is a part of is identified.

**209** An interesting question that, however, is beyond the scope of this study is whether perceptual processes in writing could also be categorized as productional processes in line with the motor theory of speech perception. If processes are applied at the perceptual stage and change the mental product, these processes could hypothetically be categorized as resembling productional natural processes. There are indeed studies in which it was shown that motor areas in the brain are activated when handwriting is being read, suggesting "embodiment of the visual perception of handwritten letters" (Longcamp, Hlushchuk & Hari 2011: 1250).

ent approach, graphemes are – at least descriptively (their psychological reality is a different matter) – defined as signs that semiotically relate material units with linguistic units. This also means that while natural phonological processes apply to units that are simultaneously material and linguistic, natural graphetic processes apply exclusively on material units, i.e. the concrete graphs produced in writing and the basic shapes they are assigned to. They do not, however, concern linguistic units. Accordingly, for speech, natural processes are much more crucial in that they determine the phoneme inventory of a language. In writing, on the other hand, natural processes can only determine the productional and perceptual categorical boundaries between basic shapes. How basic shapes that are perceived as different visual units then enter into graphematic relationships with linguistic units to form graphemes is the next step and *not* determined by natural processes. Thus, if a sequence of two basic shapes is hard to write by hand, the motoric program involved might change to alleviate the production process – which might or might not be accompanied by a change of how the basic shapes visually appear. In general, however, it is fair to say that in natural graphetic processes, basic shapes are commonly not substituted by other shapes that take part in different graphematic relationships. In speech, by contrast, changed segments might indeed have a different phonological status than the original segments, such as when /s+d/ changes to either [st] in Basque or [zd] in Spanish (cf. Hurch & Nathan 1996: 234). Concerning this very example, the question can be asked of how the alteration of one feature of a phoneme can be transferred to natural graphetic processes. Arguably, a conceptual transfer of this kind would require a consistent descriptive approach of segmenting graphs (or better basic shapes) of all of the world's scripts into respective features, an enterprise that has hitherto not been successful – not only but predominantly because of the vast visual richness exhibited by the myriad scripts of the world (but cf. Section 5.1).

An illustrative example of shapes that developed variants which are so distinct that they have become conventionalized as different allographs of one grapheme is Arabic script and its intra-inventory graphematic allography (cf. Section 2.3). From a synchronic perspective, using different position-dependent basic shapes to realize certain graphemes is conceptually comparable to the use of sentence-initial majuscules in alphabets, which was characterized as a graphematic rule above. However, similar to morphophonological rules that might have originated as natural phonological processes and to this day show remnants of a phonetic motivation, this graphematic rule in Arabic, too, started out as a natural graphetic process: The different positional allographs of graphemes, as is still visible in the shapes' appearance, are suited to the position in which they appear within a word, more specifically to the graphetic connections that need to be made on one or both sides of them. Thus, the variant shapes are motivated by the articulatory pressure to connect shapes with their surrounding shapes – a natural graphetic process (cf. also Section 7.1.2). However, what started out as automated variants has developed into a conventional rule so that one basic shape with four variants was split up into four distinct basic shapes. At this stage, then, whether these shapes that belong to one grapheme are visually similar is of secondary interest. What is more relevant is that the variants for a specific position (initial, medial, final) are adjusted to the position they appear in (with non-connecting shapes being exceptions) as well as to the graphetic connections that need to be made. In short, the synchronic visual appearance of Arabic script is still visibly determined by this natural graphetic process, even if that process has morphed into a rule.

As established above, natural graphetic processes are variable. Consequently, just like in NP, two types of lenitions and fortitions must be differentiated: On the one hand, there are those that apply or fail to apply purely on the basis of physiology, e.g. because a sequence of basic shapes is hard to produce physiologically. This type of lenitions and fortitions is non-semiotic. On the other hand, the application of processes – precisely because it is neither obligatory nor universal – can also be sociosemiotically charged. Gordian Schreiber (submitted) provides a riveting example of this. In his study of Japanese letter writing in pre-modern times (11th–13th centuries), he shows how cursivized handwriting signaled not only situations of more "informal" writing but more importantly, social hierarchies: When a person from a socially higher level addressed a person from a socially lower level, their handwriting could become very cursive, whereas when a person addressed someone from the same or a higher social level, handwriting tended to be a lot more meticulous, and the produced graphs resembled their prototypical basic shapes more closely. Schreiber terms this phenomenon *visual politeness*. In this context, the "quality" of the written product was not dependent on physiological conditions but was determined sociopragmatically. Here, the application of graphetic processes functions as a sign. This, of course, works only when both senders and addressees share the relevant communicative knowledge to interpret the appearance of writing as semiotically significant.

In a nutshell, lenitions and fortitions are not necessarily indexical of the communicative situation and its various aspects but definitely *can* be – even if this is not intended by senders or addressees. It is trivial to point out that the appearance of handwriting on a shopping list a person has written for themselves might differ from the appearance of handwriting on a birthday card intended for someone else, but this is exactly where lenitions and fortitions are semiotic in nature.

Now, when writers seek to improve the *readability* of written utterances, this involves not (only) natural graphetic processes, but natural graphematic processes. A crucial difference between these two is that graphematic processes are motivated predominantly by conscious choices. For these graphematic processes, there exists no real analog in NP. Consider a speaker who attempts to consciously enhance the acoustic quality of her speech: This conscious choice would result in (probably unconscious) fortitions and this would enhance intelligibility.[210] What, however, could be a phonological analog to a writer's conscious deci-

**210** Note that the decision to speak more clearly is conscious. By contrast, the processes that this decision evokes might not be (or do not have to be) conscious. Crucially, the sion of avoiding rare abbreviations in writing and to instead fully spell out words or to start a new paragraph to enhance the structure of a text in order to make it more readable?[211] In phonology, a possible analog could be something like natural rhetorical processes that are employed to make speech more comprehensible, examples being the conscious choice not to produce spoken sentences that are too long or the decision to make more pauses. These choices concern not the material quality of speech (volume, speech rate, etc.), but the structure of speech. The pair *legible* vs. *readable* in writing could thus be an analog of *intelligible* vs. *comprehensible* in speech.

Reconsider using or avoiding abbreviations as a natural graphematic process. From the perspective of the writer, the use of abbreviations usually represents a lenition. For the reader, on the other hand, abbreviations require (albeit often not much) extra effort to decode a written utterance. Note that since graphematics always subsumes graphetics in order to be 'visible', graphematic processes are also intricately linked to the graphetic level. For example, when someone chooses to write "e.g." instead of spelling out "for example", this might not only be more economical conceptually (if it is that at all[212]) but also graphetically since it involves less motoric effort. Vice versa, printing a word in bold when producing text on an electronic device requires extra effort for writers both graphematically, as they have to make the conscious choice of altering the structure and visual appearance of the text, and graphetically, as they have to press a button or combination of keys to actually render a portion of the text bold. Inversely, processes like this often improve both legibility and readability for readers. Less or more graphematic effort, thus, often equals less or more graphetic effort.

To sum up, there exist three types of natural processes in writing: (1a) *Physiologically conditioned unconscious graphetic processes* that change the shapes of handwritten graphs. They are determined solely by physiological limitations that writers are rarely aware of, rendering them unconsciously motivated. Secondly, there are also (1b) *sociopragmatically conditioned unconscious graphetic processes*. Crucially, although they are unconscious, they are caused by conscious choices. They include processes invoked by writers' decisions to produce texts in an aesthetically pleasing way (writing legibly by hand, choosing a typeface that is conventionally thought of as beautiful, …). At their core, these processes are still physiologically conditioned; however, whether they apply or not is determined by other

underlying motivation for these processes is different: They are sociopragmatically motivated.

**211** Something similar can of course be done in spoken language, when speakers avoid words they suspect to be unfamiliar to hearers. This, while occurring in spoken language, is not a phonological analog, however, but a lexical one that is actually independent of modality.

**212** Just because they are shorter, the use of abbreviations does not have to be more natural for writers. Indeed, the fact that they are physiologically more economical does not mean that they are automatically also less expensive cognitively. Without empirical evidence, however, these are only assumptions.

reasons, most often of sociopragmatic nature. Accordingly, graphetic features of texts can often aid in the analytical reconstruction of aspects of the writing situation such as the relationship between the writer and the addressee. Next, there are actual (2) *conscious graphetic processes* such as setting something in italics in a digital document. In this case, not only the decision to make something more legible is conscious but also the process itself. Finally, there are (3) *natural graphematic processes*. They are always conscious and necessarily go hand in hand with graphetic processes: As argued above, using an abbreviation such as "e.g.", since it must also be materialized, is always also a graphetic process.

### 4.2.3 Naturalness parameters

Nothing in the world is good or bad *an sich*.
(Vennemann 1988: 1, emphasis in original)

The semiotic-cognitive nature of NM suggests there must be various dimensions with respect to which naturalness of morphemes and words can be determined, since the semiotic relation – in Peircean terms – between a signatum and a signans can be evaluated in more than one respect. Indeed, in NM, naturalness is interpreted as a gradual, evaluative notion (cf. Dressler & Kilani-Schoch 2016: 357). However, it is not instrumentalized to rank the absolute, that is, overall naturalness of linguistic phenomena (both large phenomena such as whole languages and smaller phenomena such as individual morphemes of a language), since these, as Vennemann's introductory quote implies, are not inherently (German *an sich*, lit. 'by itself') good or bad. Instead, an assumed overall naturalness of linguistic phenomena is always the approximate sum of a bundle of values on several so-called *naturalness parameters*. A comparison of naturalness, thus, is only reasonable when made with respect to a specific parameter. And even naturalness on a given parameter is commonly not absolute but a matter of degree. There is "no general agreement on the number" of these parameters, "nor is there a fixed list of semiotic principles […] from which parameters can be deduced" (Crocco Galèas 1998: 22). This conveniently leaves the door open for new parameters to be added by a transfer of the concept of naturalness parameters to grapholinguistics.

Indeed, the semiotic naturalness parameters of NM can be operationalized for the study of the naturalness of writing. This is possible due to the fact that graphemes and graphematic units, in general, are conceptualized as signs (cf. Section 2.1) that relate a visual signans with a linguistic (phonological, morphological, …) signatum. Crucially, with respect to the signatum, it is paramount to note that there exist two types of graphemes: (i) *phonographic graphemes*, i.e. those that have a phonological unit that does not itself bear meaning as their signatum, are primary signs. They relate a visual basic shape such as |s| with a phoneme such as /s/ or a basic shape such as |り| with a phonological syllable such as /ri/. Since phonemes and syllables are themselves not signs, the graphemes representing them are signs of the primary order. By contrast, (ii) *morphographic graphemes* are graphemes of a secondary order – they are signs of signs: The basic shape |木|, for example, is in a semiotic relation with the morpheme {TREE}. This morpheme is itself a sign that relates the phonological representation /mù/ with the meaning 'tree'. Notably, in many cases, this distinction of two types of graphemes can be disregarded as it is not expected to make a difference in general graphematic analyses. The specifics of the semiotic relationship between the visual basic shape and the linguistic units (whether phonemes, syllables, or morphemes) is always of greatest importance, and this is certainly true both for phonographic and morphographic graphemes. However, in the latter, we are confronted with an additional layer of complexity as morphographic graphemes can refer not only to the morphological level itself but simultaneously to the phonological level which, given the double articulation of language, is always a necessary constituent of the morphemes.

The semiotic naturalness parameters described below can be applied to graphemes as signs, grapheme inventories as sets of signs, and, ultimately, writing systems as semiotic systems. Although it was a core assumption in NM that the features of the semiotic relationship between signans and signatum bear on cognitive processing, the question of how precisely they affect specific cognitive processes remained largely implicit. By comparison, as a more or less direct offshoot of cognitive linguistics, usage-based approaches to linguistics (cf. Ibbotson 2013) have a lot more to say about "a wide range of cognitive and social processes" (Diessel 2017) which are subsumed under the three general categories of social cognition, conceptualization, and memory and processing. It is in this context that NM might want to borrow some traits from usage-based approaches. As it is now, the ideas of NM, given its main interest in static semiotic relationships and the naturalness parameters based on them, are most relevant in assessing the linguistic fit and are in need of an extension in order to be able to also evaluate the processing and sociocultural fits of writing systems.

Like some natural processes that affect both graphetics and graphematics, some of the naturalness parameters introduced below can be applied to graphetics as well as graphematics. This is because some parameters (such as optimal shape and binarity, see below) are not concerned with the semiotic relationship between the two constitutive parts of a sign but with other features of the sign as a whole instead, and this includes aspects of materiality. For instance, that a large proportion of Chinese graphemes is characterized by a binary structure given that they consist of a semantic component and a phonological component (or a semantic component and a second semantic component, …) or that lowercase basic shapes of Roman script such as |d| can be analyzed as hierarchically complex structures with a head (the vertical stroke, |l|) that is primary and a coda (the smaller curve, |c|) that is secondary are both graphematic and graphetic or even solely graphetic applications of the parameter of binarity.

What follows is merely a short characterization of the parameters described in NM (for additional parameters and examples, see Crocco Galèas 1998) as transferred to grapholinguistics (examples can also be found in Meletis 2018:

76–80). They will be treated in-depth in the context of the linguistic (Chapter 6) and graphematic processing fits (Section 7.2).

First, we have the parameter of 1) *constructional iconicity*, more widely referred to as *diagrammatic iconicity* or simply *diagrammaticity*. The semiotic principle motivating this parameter holds that a semantically marked category should correlate with an increase of substance in the signans, e.g. as in SG *boy*, PL *boy+s* (cf. Wurzel 1994: 2592), where the category plural is more marked and the form of the plural is also marked (by the addition of *-s*). This parameter can be transferred directly to grapholinguistics to assess whether there is a diagrammatic relationship between the visual signans and the linguistic signatum. An obvious example of this is Chinese <木> *mù* 'tree' and <森林> *sēnlín* 'woods'. Here, the conceptual semantic increase in the signata of the two morphemes (a single tree vs. the woods, i.e. many trees), which are the respective signata of the graphemes, is reflected diagrammatically by the increase of material substance in the basic shapes, the graphemes' signantia. Notably, this material increase is doubly diagrammatic as it is not any arbitrary increase in graphic material but the basic shape used to represent the morpheme '(single) tree' that is reduplicated to represent 'many trees'. It is noteworthy, however, that the reduplicated shapes are adjusted in order to fit the fixed size of the segmental space in Chinese script. In short, the parameter of diagrammaticity is central in an investigation of grapholinguistic naturalness, as illustrated here by this morphographic example. Moreover, other types of iconicity are also relevant in writing: One of them is imagic iconicity, which, in writing, is known as *pictography*. Take again <木> *mù* 'tree'. This grapheme is, even after thousands of years of use and development, still partially pictographic as it visually resembles a tree. A different example is basic shapes in Korean Hangul which (roughly) depict the places of articulation of the phonemes they are in graphematic relations with (cf. Section 2.7 and Lee 2009; Kim 2011; Sung-ik 2016). In the context of the linguistic and processing fits, the different types of iconicity (especially its subtype of diagrammaticity) will be evaluated in detail.

According to the parameter of 2) *morphosemantic transparency*, a word is maximally morphosemantically transparent if its meaning is fully compositional, i.e. if it equals the sum of the meanings of its constituent morphemes. The meaning of the inflectional form *bird+s*, for example, is comprised of the meaning of *bird* plus the meaning of plurality (cf. Dressler & Kilani-Schoch 2016: 363). Concerning writing, at first glance, this parameter is also most relevant in morphographic writing systems. In Chinese graphemes of the *huíyì type*, i.e. semantic compounds (cf. Gong 2006: 45–47), two semantic components are usually combined to represent a morpheme whose meaning subsumes their two meanings compositionally. For instance, 'hand' and 'tree' are combined to form the morpheme 'to pluck, to pick'. Graphemes that are transparent in this way are maximally natural on the parameter of morphosemantic transparency. Since the parameter is interested in the nature of the composition of elements and can also be applied to phonographic graphemes, it shall be renamed *compositional transparency* in the context of a theory of writing.

3) *Morphotactic transparency* concerns complex or inflected word forms. These are considered natural if their "constituents can be perceived without opacifying obstructions" (Dressler & Kilani-Schoch 2016: 364). Accordingly, a form such as *fife+s*, the plural of *fife*, is morphotactically more transparent than *wive+s*, plural of *wife*, because, in the latter, a morphonological rule has opacified the form of the singular base. In writing, this parameter is concerned with linearity and the question of how basic shapes or even the graphemes they embody align with the linguistic units they relate to. Here, unsurprisingly, units larger than individual graphemes come to the forefront: While individual graphemes might be transparent, in the context of a larger unit such as the graphematic word, they might not be sequenced in accordance with the order of linguistic units they represent. In Thai, for example, there are so-called *misaligned* or *non-aligned vowels* as in the graphematic word <แบน>, which represents the lexeme /bɛːn/ 'flat'. Notably, the actual graphemes appear in the sequence *<ɛːbn>, i.e. the vowel grapheme <แ> precedes the consonant grapheme <บ> /b/ despite the fact that in the phonological representation of the word, the vowel phoneme follows the consonant phoneme (cf. Winskel 2009: 22). In other words, in this example, the parts of the signans of the graphematic word are not aligned with the parts of its signatum. In NM terms, this can be interpreted as an unnatural "opacifying obstruction" (Dressler & Kilani-Schoch 2016: 364). Since in the context of writing, this parameter is not necessarily related to morphology, it shall be renamed *positional transparency*.

The parameters of *uniformity* and *transparency* are subsumed under the heading of 4) *biuniqueness*. If the relationship between the signans and signatum in a sign is both uniform and transparent, meaning "one and the same form has always the same meaning and, vice versa, one and the same meaning is morphologically expressed only by one form" (Dressler & Kilani-Schoch 2016: 366), it is said to be biunique. In writing, 4a) *uniformity* and 4b) *transparency* can be used to assess the shallowness or opacity of the graphematic and orthographic modules. Are linguistic units uniformly represented by basic shapes? This is central in spelling, i.e. active production processes. Inversely, do basic shapes represent linguistic units transparently? This is relevant for reading, i.e. perception processes.

The parameter of 5) *optimal shape* is concerned mainly with the length of signantia. As words are regarded as primary signs in NM (cf. Dressler 1987: 168), this particularly concerns the length of (complex) words. Semiotically, this parameter is motivated by the fact that signs are expected to be distinctive and salient, and length is assumed to play a crucial role in perception and retrieval processes. This parameter is the first in this list that does not necessarily concern the (descriptively assessed) semiotic relationship between signatum and signans but the complexity of either the whole sign or only the signans of the sign. It is, in the approach presented here, treated exclusively as a matter of the processing fit. If evaluated graphematically, optimal shape concerns the shape of the whole sign, while graphetically, it assesses only the shape of the signans. Graphematically, the most central question is: What is the size of graphematic units, most reasonably interpreted as *length* in terms of the number of units, that is most natural for cognitive processing? As mentioned above, in NM, words were regarded as primary signs (cf. Dressler 1987: 168), i.e. signs of primary importance for processing. A simple transfer of this postulate to writing is problematic since there is no coherent conception of a graphematic word across writing systems (cf. Section 2.5). Notably, if the relevant unit is not a graphematic word that was defined independently, i.e. with features inherent in the writing system, but instead a written unit that represents morphosyntactic words, then the question of how these latter morphosyntactic words are represented in the writing systems of the world and whether they are, for some reason, more central for processing than other graphematic units (such as the grapheme) constitutes an intriguing question. It has actually been asked rather often in the form of the question of whether we process individual units (e.g. letters) or whole words in the reading process. Graphetically, i.e. when only the signans is of concern, the parameter of optimal shape invokes the concept of visual, or more generally, graphetic complexity. How much complexity is natural for the processing (including both articulation and perception) of a basic shape or a sequence of basic shapes?

Some possible examples of a grapholinguistic interpretation of the parameter of 6) *binarity*, holding that binary structures are preferred to ternary or *n*-ary ones, were already given above in the context of other parameters. They include the binary structure of many Chinese graphemes or the hierarchic binary structure of basic shapes in Roman script. In the following, this parameter will not be treated separately as its exact value for a theory of writing is not straightforwardly clear. However, its relevance might be discovered and re-evaluated in future developments of a theory of writing.

The next parameter, 7) *indexicality*, deals with the temporal (in speech) and spatial (in writing) distance or proximity between the different parts of a sign and with indexical relations between them. Here, a possible graphematic question is how the different subsegmental components of Chinese graphemes are positioned spatially within the segmental space. A more global question concerns higher levels, i.e. textual organization. In Natural Textlinguistics, a minor subbranch of Naturalness Theory (cf. Dressler 1989), indexicality is interpreted intratextually and evaluates the relation between indexes and indexed signs in texts. This is especially relevant for complex texts such as this book that consist of different classes of elements – main text, headlines, footnotes, figures, etc. There are complex indexical relations between these elements, and questions of layout and spatial arrangement significantly affect the naturalness of such texts for processing.

Like optimal shape, 8) *figure—ground* is a parameter of perceptual nature. Arguably the most salient distinction of figure and ground in writing is the one between text (figure) vs. non-text (ground), i.e. between segmental spaces that are occupied by graphic material vs. empty spaces. Empty spaces are indeed crucial to make visible different types of written units in the first place, and this is expected to have a major influence on the processing of writing. Several other applications of the parameter are imaginable, one of which is how uppercase basic shapes in scripts with a case distinction are visually more salient than lowercase basic shapes. Considering abjads and abugidas, are the functionally and visually secondary vowel graphemes (which are often even referred to as 'diacritics') the grounds to the more salient consonant graphemes, which are the figures? In general, concerning writing system typology, why does there appear to be a primacy of consonants, i.e. why are consonants much more often made out to be the salient figures to vowels as less salient grounds?

This leads to a question that is unique to a prospective theory of writing since it deals with writing as a secondary semiotic system: Which type of unit of the primary system, i.e. language, is the most natural candidate to be represented by units of writing: phonemes, syllables, morphemes? Crucially, this question must be evaluated separately for the linguistic and processing fits: The respective parameters are termed *unit of representation* and *unit of processing*. Differences between them underline clearly that a distinction between these two fits is not only reasonable but theoretically necessary. With respect to the linguistic fit, when assessing the naturalness of different types of linguistic units, the structure of a given language must be considered; here, its type (agglutinating, isolating, etc.) is already a strong indicator. As already mentioned above, an alphabetic writing system would not suit the varieties of Chinese due to their system-specific features, primarily their morphosyllabicity and the lack of inflections. These features affect what is regarded as the most natural unit of representation, which, in Chinese, appears to be the morpheme. Notably, this parameter interacts fundamentally with other variables such as the size of the grapheme inventory: A morphographic writing system, for example, will always have more units than a syllabographic or segmental system. Crucially, what is the most natural unit for processing must be evaluated in the next step. Based on a variety of historical and experimental evidence, it has been postulated that the phonological syllable is universally the most natural unit of processing (cf. Daniels 2018). This claim will be discussed in detail in Section 7.2.1.

A given graphetic or graphematic unit can never simultaneously exhibit the most natural values on all of the naturalness parameters since these are in conflict with each other, a phenomenon referred to as *naturalness conflict*. Naturalness conflicts are particularly useful in showcasing how different scripts or writing systems or, at a typological level, types of scripts or writing systems deal with the same kinds of challenges in different ways.

Several more parameters were postulated in NM, some of which are treated systematically in Crocco Galèas (1998). A question that I previously raised and affirmed was whether it is to be expected that more parameters will be found in the course of a transfer of the concept to grapholinguistics (cf. Meletis 2018). However, whether this claim will turn out to be true depends on the definition of *parameter*. If parameters are to be understood exclusively as facets of the relationship between signans and signatum, then even some of the parameters listed above technically do not count (such as optimal shape or unit of operation) as they are not concerned with this relationship but with other features of the entire sign or even inventory of signs. For this reason, I opt for a broader reading of *parameter* that allows for the inclusion of other parameters such as *unit of representation* and *unit of processing* (cf. Sections 6.1 and 7.2.1). Given that NM focuses on the semiotics of morphological structure, the naturalness parameters that have been established in this context are only peripherally concerned with materiality. I argue that there are indeed also material naturalness parameters. Above, it was explained that there are graphematic (i.e. linguistically functional) natural processes in writing; similarly, there are not only graphematic naturalness parameters but also graphetic, i.e. material ones.

To recap, NM treats synchronic linguistic structures as signs and investigates the facets of the semiotic relationship between their constituents. One type of external evidence used to evaluate which types of semiotic relationships are more/less natural is language acquisition (see below). Here, the order in which children acquire features of language is taken to reflect their degrees of naturalness; additionally, mistakes in the production of children also point to the naturalness of different morphological features. This is a static, product-oriented approach. Notably, one can also adopt a more dynamic perspective: What if the mistakes that children produce during language acquisition were to be interpreted as the outputs of so-called natural morphological processes? When children overgeneralize, i.e. produce a form in a way that is regular in a given paradigm although the correct form is actually irregular, such as German **geschmeißt* for correct irregular *geschmissen* 'thrown' as an incorrect derivation from *schmeißen* 'to throw',[213] this might be conceivable as a natural morphological process. In fact, analogy of morphological structures is listed by natural morphologists as external evidence. Arguably, analogy is a natural morphological process both ontogenetically, in children's language acquisition, and phylogenetically, in the diachronic development of languages (i.e. in the course of language change). That way, the static semiotic structures studied by NM can be interpreted as the results of dynamic natural processes. For an analogous example in writing, take a spelling error in German, *<foll> instead of orthographically correct <voll> 'full'. Here, the normatively 'wrong' grapheme is chosen probably because the basic shape |f| is in a much more regular and straightforward relationship with the phoneme /f/ than |v| (cf. Balestra, Appelt & Neef 2014), and the fact that <f> is more natural with respect to the naturalness parameter of transparency leads to errors of this type. That the ensuing output – much like the example taken from language acquisition above – is (orthographically) incorrect is not relevant since the normative orthographic level actually needs to be disregarded in an investigation of the actual system behind it. Indeed, *<foll>, although orthographically incorrect, is certainly part of the graphematic solution space of the German word *voll* 'full'. This mistake of producing <f> instead of <v> can be modeled as a natural graphematic process caused by the fact that no grapheme uniformly corresponds with the phoneme

**213** I thank my nephew for giving me this example when he was 4 years old.

/f/ in the writing system of German – and as postulated by NM, semiotic structure affects cognitive processing.

Inversely, naturalness parameters, characterized by being static, can be transferred to NP. The reasons that natural phonological processes even occur can be explained from a static, product-oriented perspective. For instance, the reason for the assimilation of one phoneme in the sequence /s+d/ to either [zd] (in Spanish) or [st] (in Basque) could be stated in terms of parameters. While it is not my intention to describe how this can be done in detail, I want to stress that in Naturalness Theory, every phenomenon can be analyzed from both a static and a dynamic perspective. Thus, when children invert the basic shape |J| during the first stages of literacy acquisition, this can be treated as a natural process caused by the cognitive difficulty of processing an exceptional left-faced shape in a system of otherwise rightwards-oriented shapes which also happen to be oriented in the dextrograde writing direction. The same difficulty can be formulated in terms of a parameter: Orientation towards the left is less natural in Roman script than orientation towards the right. What remains to be investigated is the level at which the naturalness of this parameter is located (the system-independent, typological, or system-dependent level). Only a comparison with similar phenomena in other scripts shows that it is not the rightward orientation or the orientation in the same direction as the statistical majority of basic shapes in a script but instead the orientation in the direction of writing that is most natural for children at this stage – and likely also in more proficient stages after that (cf. Section 7.1).

Building on this brief introduction, Chapter 6 is devoted to an in-depth analysis of how the naturalness parameters of NM and specific additional parameters such as unit of representation can contribute to an evaluation of the linguistic fit of writing systems. There, many examples from various systems will be discussed. Section 7.2 will then shift the focus to the question of how the same parameters affect processing.

### 4.2.4 Methodology and evidence

Naturalness Theory does not have a clear-cut methodology. However, one methodological aspect is undeniably central to the approach: the consideration of external evidence, which is why in this section, I will discuss how different types of evidence can be used to uncover natural features in writing systems. The types of evidence that were most crucial in NP and NM are also relevant for a grapholinguistic theory: (1) acquisition, (2) mistakes/errors, (3) disorders, and (4) change (cf. Watt 1975: 297).

In the context of (1) the acquisition of a first writing system, which, by analogy with L1 acquisition, can be termed *L1WS acquisition* (cf. Cook & Bassetti 2005), the central question is: which features and structures of writing do children acquire first or relatively early when they learn to read and write? Crucially, this question must be treated differently than the question of the sequence in which linguistic elements and structures – phonemes, morphemes, syntactic structures, etc. – are acquired in the course of L1 acquisition. Language is – in the prototypical case – "inborn" in the sense that it seems to require no extreme and straining effort for children to acquire (this is not to be misread as a generativist claim to innateness). Writing differs from language in that it is taught and learned through instruction which, in turn, is determined not only by individual teachers but also top-down in the form of curricula prescribed by governments. Thus, the ontogeny of language and writing is fundamentally distinct. Since it is determined externally, the order in which children acquire grapholinguistic elements and structures cannot be readily regarded as particularly insightful. However, 'acquiring' and 'mastering' are two different things, and the order in which children *master* certain features of writing after being familiarized with them through instruction is indeed revelatory.

It must be noted that children can acquire rudimentary knowledge about writing before they are formally instructed. For example, in Anbar's (1986) study, six preschool children from different backgrounds and with varying IQs learned to read without instruction. Notably, their individual processes of reading acquisition exhibit significantly similar patterns. Anbar believes this finding indicates "a *natural* process taking place in the reading development of preschool children who grow up in a particularly literate environment". *Natural*, here, is supposed to mean "that this process is neither directed nor guided from outside the child by the parents or by some standard reading method, but rather develops within each child as a result of something internal to him or her" (Anbar 1986: 78, my emphasis). In other words, *natural* is equated with 'acquired without instruction'. In Section 7.1, in the context of the graphetic processing fit, I will discuss more findings from studies on emergent literacy that show what children learn before being instructed.

The next type of evidence is closely tied to literacy acquisition: (2) the analysis of mistakes, sometimes referred to as *mistake and error linguistics.* It focuses not only on mistakes made in acquisition but on mistakes of writing and reading in general. Thus, studies are conducted both on mistakes and errors of children and those of healthy as well as impaired adults. Note that following Corder (1967), I distinguish conceptually and terminologically between *error* and *mistake*: Errors occur when someone produces something incorrectly because they lack the knowledge necessary to produce it correctly (or have erroneous knowledge). Thus, errors are both unintentional and non-corrigible. Mistakes, on the other hand, occur when someone produces something incorrectly despite knowing the correct form, i.e. "the form […] selected was not the one intended" (cf. James [1998] 2013: 77). Consequently, mistakes are often corrected by the person who made them. The most obvious mistake in writing nowadays is the so-called *typo*, i.e. when someone accidentally switches two letters when typing on a keyboard. In short, errors are phenomena of *competence*, mistakes matters of *performance*. Crucially, a broader definition of *error* relevant in grapholinguistics also includes conscious choices that deviate from the orthographic standard: Examples are the omission of punctuation marks or, especially in German (but also other alphabets with a case distinction), the neglect of rules pertaining to capitalization.

Further evidence that can reveal natural features in writing comes from the analysis of how aphasics or, more generally, people suffering from (3) disorders of reading and written expression deal with processes involved in reading and writing (cf. Reitz 1994; Gregg 1995; McCardle et. al 2011). The broadest naturalist hypothesis with respect to this type of evidence is that stable elements, i.e. those that are retained (longer) despite various impairments of processing, are more natural than those affected (earlier) by them.

The final central type of evidence I want to mention here is (4) change. For a functional theory of grapholinguistics, the development of writing systems is "relevant to understanding the differential cognitive demands of language and reading" (Perfetti & Harris 2013: 297). As Naturalness Theory predicts that systems tend to eliminate unnatural features, changes in scripts and writing systems will likely point to unnatural features – and, in this process, will also uncover natural features. An aspect of Naturalness Theory that is often misunderstood and, in turn, adamantly criticized is its supposed teleology. It is said to claim that systems change to become more natural and to finally arrive at an "ideal" state. This, however, is logically impossible. Change – and this is true both for language and for scripts and writing systems, all of which are, crucially, systems – "is local and not global because of goal conflicts which characterize all functional systems; [...] owing to the tendency of each component of grammar to increase its own naturalness, markedness reduction on one level usually brings about markedness increase on another" (Bertacca 2002: 9). Take the basic shapes of Arabic script: they developed in a way that made them increasingly similar, which resulted in a perceptual decrease of distinctiveness that is unnatural for the reading process. The reason for this might be that production was primary over a critical period of time, and the priority was to minimize the number of motoric programs that scribes must memorize. This can – but does not have to – lead to visually similar basic shapes (cf. Section 7.1.2).

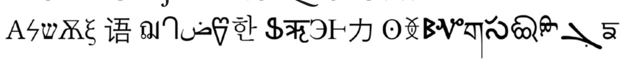
Finally, it is noteworthy that writing enjoys a special status: While the above-mentioned types of evidence are relevant to uncover natural features in writing, writing itself also serves as external evidence in the investigation of phonology and morphology (cf. Zwicky 1973: 88). Specifically, since writing always represents an analysis of the underlying language system, the question of which features of language are reflected in writing and which are not can be utilized as evidence for what is natural in language.

# 5 Systematic fit

As established above, a writing system consists of separate modules (cf. the introduction to Part II). These modules – the graphetic, graphematic, and orthographic modules – are systems in and of themselves. This means that for each of them we can evaluate, in a nutshell, the systematicity of their units – basic shapes, graphemes (and larger graphematic units), and orthographic rules. ‘Systematicity’, here, means both how systematic the units of a system are designed and whether the relationships between them are systematic. Do the units share features that are consequently characteristic of the system? Are these features evenly spread throughout the system or are there many (unexplainable) idiosyncrasies? The systematic fit is both purely descriptive, i.e. independent of the use of these systems (with its processing and sociocultural aspects), and self-contained. The graphetic systematic fit, for example, is concerned purely with the systematicity of this module and not with the question of how systematically it interacts with other modules (such as the linguistic fit, which evaluates how well the graphematic module fits the underlying language system). In the following, I will discuss what is necessary to evaluate the graphetic systematic fit (Section 5.1) and give an illustrative example of how the orthographic module can exhibit an unnatural systematic fit (Section 5.2).

## 5.1 Graphetics

ABCDEFGHIJKLMNOPQRSTUVWXYZ



Above, you see two sets of 26 basic shapes each. One of them is a system, the other is not. The reason for that is that one of them consists of basic shapes that have a number of features in common (because they stem from the same script), while the other is a random selection of basic shapes taken from different scripts. Interestingly, it is likely that even non-literate people unfamiliar with Roman script would choose the first set if asked which of them is more systematic. Indeed, studies have found that children who are not yet literate reject as writing shapes that are visually dissimilar from the shapes of their own script (cf. Section 7.1.1). Aspects that account for the perception of shapes as a coherent system are subsumed under the notion of *systematic fit*. This designation is based on the fact that scripts are visual systems, sets of units that share features and are related to one another. All scripts – e.g. Roman script, Chinese script, the kana scripts of Japanese – are, from a productional point of view, graphomotoric and, from a descriptive and perceptual point of view, visual inventories that can be studied independently of the linguistic structures they are related to when they are used for specific writing systems.

If studied purely for themselves, the material properties of scripts come to the forefront, and more specifically, the question of how these properties are distributed throughout scripts. To give an example: while most of the uppercase basic shapes in Roman script are either vertically symmetrical (such as |M|) or face rightward (such as |R|), one outlier faces leftwards: |J| (see for details below). Outliers such as this are characterized by the fact that their features do not conform to the statistically frequent features exhibited by the majority of a script's basic shapes. Primus & Wagner (2013) call these outliers *non-canonical* whereas basic shapes that do exhibit a script's typical features are *canonical*. Aside from |J|, the uppercase basic shapes of Roman script are entirely systematic with respect to the orientation of the coda (if they have a coda, that is). This type of systematicity in a script has not only an effect on how it is processed (cf. Section 7.1) but, crucially, is at the same time also likely a product of how scripts are processed. Even if one analyzes scripts descriptively – as is done here – it must be kept in mind that they are not systems that have emerged independently – they are man-made systems. In this vein, Watt (1979: 31) claims that the systematicity of scripts can "be traced to that property of the human mind, ultimately the human brain, that forms systems in the first place". In other words, it is the human brain that not only enabled the invention of scripts in the first place but that thereafter also led to their increasing systematization. As the forces involved in the human-led change of scripts are crucially governed by human processing needs, they will be discussed in the context of the processing fit in Chapter 7. This chapter is focused rather on identifying and describing the inherent properties that render a script a system.

To explain the systematic fit, I want to use and modify an example originally introduced by Watt (1983a). In order to describe a system consisting of four basic shapes, two features with two feature values each are required, i.e. two binary features. Watt (1983a: 384) calculates this with the formula $V^F=N$, in which V is the number of feature values, F is the number of features and N is the number of basic shapes that can be generated by the features and their values. If, as in Watt's example, two binary features ([±cardinal] and [±top]) refer to the codas of basic shapes, this generates a total of four basic shapes that all differ with respect to their codas (cf. Table 7). In this case, all possible shapes that the binary features can generate are exhausted, meaning there are no additional basic shapes that conform to the features of this invented script and are not already part of this four-shape set. Thus, as will become relevant below, this script has no so-called *systematic gaps*, and with that, there is no redundancy in the system (for the relevance of redundancy, cf. Section 7.1.3). Furthermore, there are no outliers in this script (such as |J| in Roman script).

Table 7. Complete system

| |  |  |  |  |
|---|---|---|---|---|
| ± cardinal | + | + | – | – |
| ± top | + | – | + | – |

Table 8. Complete system consisting of an incomplete system (the script) and systematic gaps

| | complete system | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | incomplete system | | | | | systematic gaps | | |
| |  |  |  |  |  |  |  |  |
| ± cardinal | + | + | – | – | – | + | + | – |
| ± top | + | – | + | – | – | + | – | + |
| ± right-oriented | + | + | + | + | – | – | – | – |

In Table 8, an additional, fifth basic shape has been added to the system. This shape deviates visibly from the others with respect to the orientation of its coda. Like |J|, it is an exception in its script. According to Watt's formula, for the description of this inventory of five basic shapes (the inventory titled 'incomplete system' on the left), two binary features do not suffice anymore. With three features, the formula changes to $2^3=8$. This, now, provides not only enough features to adequately describe the five basic shapes but has additional important consequences: it produces three more well-formed basic shapes (positioned on the right), which, however, are not part of the system. As such, they represent systematic gaps. A system without such gaps, i.e. a system in which all possible well-formed basic shapes are actually units of the script, is a *complete system*,[214] whereas a script that has such gaps is an *incomplete system*. This explains why ||, just like |J|, remains an outlier: it is the sole basic shape in this script that exhibits the feature value [–right-oriented]. All the other possible shapes with the same feature value are systematic gaps.

214 It is a complete system in that it is maximally systematic. Note, however, that this does not necessarily make it a *closed system*: new units can potentially be added to such a system. This turns it into an incomplete system again as the added units cannot be described solely with the existing systematic features of the system, since these are already fully realized. This implies that basic shapes that are added to a complete system inevitably introduce new features. If, again, all of the feature combinations made possible by these newly added features are fully realized by the range of newly added basic shapes and there are no systematic gaps again, the system has become complete once again. Accordingly, no script is ever "closed" in the sense of closed systems. Scripts are always open systems, even if in practice, new units are rarely – if ever – added to scripts.

This makes |ᐯ| not just the only shape that deviates from the others with respect to orientation but simultaneously also the only shape for which this feature is even of relevance. Now, if the three systematic gaps were also units of the inventory, the system would be complete, and this would arguably equal the most natural systematic fit. No redundancy would exist in the system and all features would be realized to their fullest potential. This also means that the graphetic solution spaces for the individual basic shapes are very restrained in that deviations easily lead to the production of another existing basic shape of the script; this would not be the case (to such a degree) in a script that has systematic gaps.[215] Thus, a script with a good systematic fit is not necessarily a script with a good processing fit. Examples of scripts that 'at first glance' appear rather systematic are the Cree script and Korean Hangul (cf. Figure 27 and Figure 28).

As will be shown in the discussion of the processing fit below, a script that exhibits the best possible systematic fit is deleterious for the cognitive process of misremembering. To elaborate: If writers forger the value of one feature of a basic shape or one of its elementary forms (such as the orientation of the coda in |d|), they likely substitute it with the value that predominates statistically in the script (e.g. [+right-oriented]) or is more suited to processing for other reasons. In the case of misremembering the orientation of the coda of |d|, this process results in |b|, a different basic shape that actually exists in the script and – in most (if not all) writing systems using Roman script – is part of a different grapheme, i.e. related to a different phoneme. By contrast, if there were a degree of redundancy in the system, misremembering could potentially lead to basic shapes that are systematic but not part of the system because they are systematic gaps. Producing them would still count as a mistake but, crucially, it would not evoke a wrong grapheme; thus, the mistake would be purely graphetic (such as mirroring |J|) rather than graphematic (such as mirroring |d|).

The fact that redundancy is crucial for the processing fit of scripts should be kept in mind for the (rare) instances in which new scripts are created from scratch. Note that redundancy, however, should not be mistaken for *exceptionality*: the feature value [–right-oriented] for the coda of |J| in Roman script is not redundant but an exception. In order to count as redundant, a few more of the twenty-something basic shapes of the script would need to be oriented toward the left – not just one. This would make the feature value a salient part of the system while still leaving some systematic gaps. Thus, the distinction between exceptionality and redundancy is gradual rather than absolute.

215 Consider, once again, the example of |T| and |Γ|. In Roman script, the graphetic solution space for |T| is larger since |Γ| is not a unit of the script but instead a systematic gap. By comparison, in Greek script, both basic shapes exist, which means the graphetic solution spaces for both are constrained to avoid misidentification (cf. Section 1.2.1).

FIGURE 27: Selection of basic shapes from Hangul, from http://www.decodeunicode.org/en/u+11FF (April 11th, 2020)

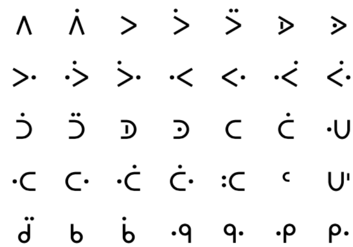


FIGURE 28: Selection of basic shapes from Cree script, from http://www.decodeunicode.org/en/u+14FF (April 11th, 2020)

A phenomenon that the systematic fit can help explain (or, in a practical sense, even guide) is that shapes that are newly added to an existing script are made to "fit" that script. In most cases, such additions are likely not designed with the systematic fit (as defined here) in mind, but they are arguably still constrained by it. For example, when the uppercase version of |ß| was added to the German writing system by the *Council for German Orthography* in 2017, it was designed to be (1) similar to the existing lowercase version that it is based on and paired with as well as (2) similar to the other uppercase basic shapes of Roman script. The result was |ẞ| (for a discussion of the criticism of this shape, cf. Section 8.3). In the future, an *a priori* analysis of the systematic fit of the script in question can guide the design of new basic shapes for it.

Notably, the systematic fit is not concerned with evaluating the naturalness of individual features such as the above-mentioned [±right-oriented]. Indeed, its descriptive methodology offers no heuristics to decide which features are descriptively more natural than others. Whether the basic shapes in a script are orientated towards the right or towards the left, thus, is not of interest for the systematic fit (whereas it is for the processing fit). In short, thus, the systematic fit deals merely with the featural (in this case this means visual) coherence of a script's units, their "degree of systematicity" (Watt 1983a). Accordingly, what must be achieved before the systematic fit can be evaluated is a description of features that provides answers to the following questions: What exactly are features of basic shapes and how can they be described?

The visual segments of basic shapes, i.e. the smaller subsegments they are composed of – lines, curves, dots – are termed *elementary forms* (cf. Section 1.2.1; for an extensive review of attempts at arriving at graphetic elementary forms and features, cf. Meletis 2015: 50–79). It is of paramount relevance for the evaluation of the systematic fit to acknowledge that these elementary forms are not themselves features but instead consist of features. The former was assumed by, among others, Althaus (1973), who, for instance, conceived of the vertical stroke and the curved stroke in |P| as two features (instead of two elementary forms). An analysis of this kind leads to problems in the featural description of basic shapes, which was most vocally expressed by Watt (1975: 303–323). He argued that two types of underdifferentiated featural analyses are ultimately doomed to fail: (1) analyses such as Althaus' that equate elementary forms with features, and (2) analyses such as Gibson et al.'s (1963) that assign visual features such as [+curved] to entire basic shapes rather than to elementary forms (and thus skipping the process of segmenting basic shapes altogether). The first of these analyses is not so much inherently inaccurate as it is uneconomic: assuming that |I| and |—| are different features neglects the fact that they are, essentially, the same elementary form, a straight stroke. They differ only in orientation, and it is this fact that should be conceptualized as a feature. The elementary forms |C| (itself also an independent basic shape) and the left-facing coda in |q|, i.e. |c|, are also instantiations of the same elementary form. They differ in size and, related to that, in the position they occupy in the segmental space (cf. Section 1.2.1). The second analysis is also underspecified: assigning features such as [+curved] to entire basic shapes such as |P| is problematic since it is evidently not the whole basic shape that exhibits this feature but only one of its components. Thus, it is fatal if such an analysis fails to offer additional tools to further specify which part of a basic shape a feature applies to. In an attempt to eliminate the shortcomings of these two types of analyses, Watt (1980: 8) uses linguistic levels as an analogy to describe how he interprets letters (or more generally, basic shapes) as analyzable units:

> Letters are morphemes because the units of the next level coöccur in sequence, which in turn is so because any solution that directly factors letters in simultaneously-coöccurring units, or features, suffers severe flaws.

The resulting assumption of the more fine-grained hierarchy of *basic shape > elementary form > feature*, unfortunately, does not solve all the remaining problems of a featural description of basic shapes and scripts. The following questions must be answered individually for every script: which are the relevant elementary forms of a script's shapes and which are their relevant features? Unlike listing the set of basic shapes of a script, accumulating elementary forms and features is not trivial.

To give an example of a featural analysis that can serve as a model for future analyses, Watt proposed several features for the uppercase basic shapes of Roman script. In the two first contributions to an article series titled *What is the proper characterization of the alphabet?*, he investigated the different features that are relevant in production (*composition*, as he calls it, cf. Watt 1980) on the one hand and in perception on the other (cf. Watt 1981). This is based on his assumption that there exist two separate competence grammars for production and perception: a *kinemic* grammar and a *phanemic* grammar. In line with his arguing, I believe perceptual features are more relevant for a descriptive account of features since description is commonly based on the descriptor's perception of the static visual datum rather than dynamic articulation processes that were involved in producing it.

The elementary forms of basic shapes, which Watt calls *kinemes* when they are viewed from a productional point of view and *phanemes* when they are treated from a perceptual perspective, exhibit the following perceptual features: [±VRTCL] 'vertical', [±HRZTL] 'horizontal', [±TRACE] 'trace', [±FLNTH] 'full length', and [±CNCVE] 'concave'. These perceptual features are all used to specify individual elementary forms, i.e. lines or strokes. [±VRTCL] gives information on whether a stroke is vertical, as in |I|. [±HRZTL] is the feature of horizontality, as evidenced by the middle stroke in |H| that connects the two vertical strokes. Being neither horizontal nor vertical, diagonal strokes are [−HRZTL] and [−VRTCL]. [±FLNTH] informs about the relative length of a stroke, i.e. whether it is, from the perspective of the segmental space, of full length, such as the vertical stroke in |R|, or not of full length, such as the two diagonal strokes in |K|. [±CNCVE] is crucial as it indicates whether a stroke is curved or not: an example of a coda exhibiting the feature [+CNCVE] is the coda of |P|. The purpose of the perceptual feature [±TRACE] is not straightforward. It is the perceptual analog to [±TRCE], the productional feature that Watt (1980) had assumed for kinemes and that distinguishes strokes that are actually written, i.e. leave a graphic trace, from strokes that are only made in the air when the writing instrument is lifted and moved to the next starting point where writing continues. These strokes in the air exhibit the feature [−TRCE]. While as a productional feature, the inclusion of [±TRCE] makes sense, it is not clear if it is of value as a descriptive (visual) feature of phanemes.

As Watt's analysis is based on the subset of Roman majuscules used in the English writing system, he assumes two more features: [±HSMTR] for horizontal symmetry and [±VSMTR] for vertical symmetry. By means of the features listed above, the two phanemes in |P| can be described as follows: the hasta is [+VRTCL], [−HRZTL], [+FLNTH], and [−CNCVE], while the coda is [+VRTCL], [−HRZTL], [−FLNTH], and [+CNCVE]. This featural analysis leaves open a slate of serious questions: is the assignment of these features really absolute in each case? Does the number of features suffice for an unambiguous description of |P|, i.e. can it distinguish |P| from other basic shapes? The answer to both of those questions is no.

To be fair, Watt's (1981) analysis is much more complex than outlined here: Watt provides not only a list of features for each basic shape but also a list of rules of how these features are applied and spatially ordered within the segmental space to result in a given well-formed basic shape. In the context of this full analysis, the above-mentioned features might indeed suffice to distinguish |P| from other Roman uppercase basic shapes. However, without Watt's (highly) complex and specific rules, or if the set of uppercase basic shapes were extended to include corresponding lowercase basic shapes, for example, the two phanemes described by the featural configuration given above could also be joined to form

|p|, |b|, |q| or |d|, all of which conform to the description of a horizontal full-length stroke and a curved half-length stroke. This goes to show that what is central for a featural description of basic shapes is information on spatial relations and the arrangement of elementary forms. How are the two elementary forms specified by the mentioned features positioned in the segmental space? Are they connected, and if so, where are they connected and how (with an acute angle, a cross, etc.)?

There are several suggestions on how spatial and topological information of this kind could be formalized. One of them will be discussed here (along with suggestions for its modification), focusing on how its core ideas could be used to establish a more universal methodology for a description of features in scripts, which is a prerequisite of evaluating their systematic fit. The proposal in question is the one made by Althaus (1973) that was already mentioned above. Althaus equates elementary forms with features and assumes twelve features. Where these features are located with respect to the vertical axis of the segmental space and over how much of the segmental space they extend is described by means of seven subspaces. These subspaces are based on the vertical subdivision of the segmental and linear spaces in the high, central, and low spaces. Four of them are 'simple', i.e. extend over only one subspace, and three of them are complex, i.e. extend over more than one subspace (cf. Figure 29). Notably, Althaus (1973) works with a more fine-grained four-space schema in which the central space is further subdivided into two spaces (in the figure, these are spaces 2 and 3, while space 5 combines the two) (cf. also Section 1.2.1). With the help of this spatial description, basic shapes consisting of only one elementary form such as |C| and |c| can be distinguished: the former occupies space 6 and the latter space 5.

In a formalization, the elementary forms are written as digits and the spaces they occupy as superscript digits. The hasta of |P| is formalized as $1^6$, the coda as $7^1$. The only information that remains unspecified is the order (or spatial sequence) of elementary forms in basic shapes that consist of more than one elementary form. Althaus suggests an arrow: $1^6 \leftarrow 7^1$ thus means that the coda ($7^1$) is located to the right of the head ($1^6$) (cf. Althaus 1973: 108).


Figure 29. Classification of elementary forms ("features") and spaces they occupy, from Althaus (1973: 108)

A problem inherent in the description of basic shapes is that the elementary forms that they are composed of cooccur. In other words, they all exist at the same time and they are arranged in space in complex ways. Granted, for a basic shape such as |P|, the description of the spatial relationship between the two elementary forms might not be complex, but for a shape such as |語|, it arguably is. In his proposed modification, Garbe (2000: 1769–1771) adopts Althaus' (1973) list of "features", i.e. elementary forms, but conceptualizes the spatial relations differently (cf. Figure 30). Using six lines (a-f), he divides the linear space into five subspaces (ab, bc, cd, de, ef). These subspaces help in formalizing the spatial arrangement of basic shapes. |T|, for example, is formalized as:

$\frac{4a}{1ae}$ Elementary form 4 is located on line a; the fraction bar signifies that it is positioned above elementary form 1 that extends from line a to line e.


FIGURE 30. A spatial analysis using six lines or five vertical subspaces, from Garbe (2000: 1769)

If elementary forms and features were conceptually kept apart, Garbe's proposal would be a promising start for a featural analysis of Roman script that takes into account spatiality. Several other attempts (cf. Bhatt 1988; Herrick 1966) are very similar in nature. Note that what is still missing in these conceptions is the description of connections. They treat all connections between elementary forms as equal, i.e. angles such as |L| and |T| but also crossings as in |X|. However, it has been suggested that these connections differ with respect to their (descriptive) complexity (cf. Altmann 2004). This makes any featural analysis that does not in any way account for these differences incomplete and underspecified.

After having highlighted that a featural description of a script's basic shapes is a prerequisite of the evaluation of the systematic fit, I am not in a position to propose or prescribe a method of description. Attaining a universal method of describing basic shapes is difficult but, I believe, not impossible. It is obvious that the vast visual variety exhibited by the world's scripts complicates this endeavor greatly. One question that should be put forward for discussion, however, is whether a formalized description such as Althaus' or Garbe's is even necessary. Instead, it could be claimed that a basic shape such as |T|, as a visual datum, is not just a basic shape but, at a meta-level, already its own description. This idea might appear odd at first, but consider that in the word de*script*ion we find the Latin word *scrībere* 'to write'. Hence, is it necessary to de-*scribe* something that is already written? I argue it is. Even if basic shapes already were their own descriptions, there would still exist a need for a vocabulary to express how different shapes are similar and how others are distinct and why it is precisely that we perceive a script as a visually coherent system (if we do, that is, as there are also incoherent systems). These questions, then, lead back to square one and underline the absolute necessity of a method of describing the featural and spatial makeup of basic shapes. This becomes even more evident when a graphetic analysis is used as a basis for a graphematic analysis, e.g. when the question is asked of whether some graphetic features are transparent, i.e. consistently correspond with linguistic values or functions (cf. the analyses in Primus 2004, 2006). In such cases, a graphematic analysis that is not based on an identification of graphetic features is set up to fail. The preceding remarks have introduced some suggestions that can help in coming up with universal or script-specific methods of description relevant in a comprehensive theory of writing.

Additionally, several variables that crucially affect the systematic fit shall not be left unmentioned: a script's (1) *size*, its (2) *complexity* (defined as the number of features necessary to describe it), and the (3) *frequency* with which individual basic shapes occur in the use of a script (for a specific writing system).

The number of basic shapes in a script is always externally constrained. In the process of the initial creation of a script, for example, it is determined by the unit of representation chosen for a writing system (e.g. phoneme, morpheme, etc., cf. Section 6.1) and is, in the next step, dependent on the size of the inventory of these units, i.e. the size of phoneme or morpheme inventories (cf. Chang, Plaut & Perfetti 2016: 67). Usually, what would be expected in this context is that creators of writing systems strive for uniformity of graphemes (cf. Section 6.5) by creating one basic shape for each linguistic unit that needs to be represented graphematically. This would result in a uniform grapheme inventory in which the number of basic shapes needed for graphemes equals the number of linguistic units of a given type (phoneme, morpheme, etc.). Note, however, that the situation just outlined is the ideal (and rare) scenario of script creation. More often, scripts are adopted, and given that new basic shapes are introduced to existing scripts relatively seldom, writing systems for which scripts are adopted must make do with the basic shapes that the scripts offer in the first place. In cases in which there are more graphemes in a writing system than available basic shapes in the adopted script, now, common strategies are the use of digraphs or the addition of diacritics to existing basic shapes. Sometimes, a single basic shape is also used for more than one grapheme, leading to a decrease of naturalness on the graphematic parameter of transparency (cf. Section 6.4). If, by contrast, the number of basic shapes is greater than the number of needed graphemes, some basic shapes of the adopted script may remain unused. In short, the size of a script is usually dependent on the language and writing system that the script was initially devised for. Because of the principle of conservatism (cf. Section 7.1.1) that is based partially on the stability and rigidity of scripts (which is itself a result of the invention of printing, keyboards, typefaces, etc.) the number of a script's basic shapes is commonly not influenced by the language that adopts the script, even if the script does not offer enough shapes for the language's linguistic units. Note that a possible practical application of the systematic fit could be that the systematic gaps (see above) that were identified can serve as (models for) new basic shapes[216] if more shapes are

216 Adding new basic shapes to a script could be more natural than modifying existing basic shapes by adding diacritics (unless these diacritics are used diagrammatically

needed. This would ensure that any basic shapes added to an existing script conform to the features of the script. However, it does not solve the (arguably more important) problem of technology, namely that new shapes are not yet encoded in Unicode and not ready for digital use.

The (2) *complexity of a script*, defined as the number of features relevant to describe its basic shapes, interacts with the above-mentioned size of a script. The more units there are in a script, the more complex it necessarily becomes, as a larger number of features becomes necessary to keep the basic shapes distinct (cf. Chang, Chen & Perfetti 2018: 438; Treiman & Kessler 2014: 163; cf. Section 7.1.3). This corresponds with Watt's (1983a) formula of $V^F=N$ mentioned at the outset of this section. Consider as an example the complexity of the basic shapes of Chinese script vs. those of Roman script. It is a quite trivial observation that in a script with thousands of basic shapes (and hundreds of subsegmental components constituting those shapes), more features are necessary than in a script with twenty-something basic shapes. Of course, even a fairly compact script, i.e. a script with only a small number of basic shapes, can exhibit "unnecessary" complexity. Thus, hypothetically, even in a script with 16 basic shapes, there could be many more than four constitutive binary features (suggested by the formula $4^2=16$). That being said, the formula $V^F=N$ is only a theoretical instrument to assess how many features are *minimally* required for a given number of basic shapes in a script. This does not mean that scripts must (or do) conform to this number. What can be postulated as a more or less rough correlation, however, is that the larger a script is, the more featurally complex the individual basic shapes must become to ensure distinctiveness. Note that the quantifiable complexity of a basic shape can also be assessed 'unsystematically', i.e. without reference to features, simply by classifying and counting the elementary forms that occur in basic shapes (as dots, straight strokes, curved strokes) as well as the types of (non-)connections between them (continuous contacts, crisp contacts, crossings, cf. the proposal by Altmann 2004).

The third variable worth mentioning is the (3) *frequency* with which individual basic shapes of a script occur in the actual use of a given writing system. This frequency is determined top-down by the frequency of use of graphemes, which, in turn, is determined by the frequency with which the linguistic units that the graphemes relate to occur in the language. Basic shapes that are statistically frequently used, i.e. produced and perceived, are shaped by an accommodation to human processing needs. For example, they become easier to write, although this process of productional facilitation must be counterbalanced by the perceptual need for sufficient distinctiveness (cf. Section 7.1). A converse hypothesis informed by usage-based approaches is that frequently used basic shapes can 'afford' to be more complex (in the sense of consisting of a greater number of elementary forms) as they are more stable in users' memories. This mirrors morphological suppletion

and transparently to represent the same linguistic feature in a number of graphemes, cf. Section 6.2) since segmental basic shapes are arguably more salient than subsegmental diacritics.

and its interaction with the mental lexicon (cf. Hippisley et al. 2004). Quantitative analyses of writing systems are scarce, but studies suggest that rarely occurring basic shapes are more complex, which speaks for the first hypothesis. Specifically, for the Chinese script, it was found that basic shapes consisting of a greater number of strokes (cf. Yu 2001; Shu et al. 2003) as well as basic shapes consisting of a greater number of components (cf. Bohn 1998) are those that are more rarely used.

## 5.2 Orthography

An example of an orthographic reform that highlights how meddling with (the systematic fit of) the orthographic module can unnaturally influence the systematicity of the underlying graphematic module concerns Chinese, specifically the changes to Chinese graphemes that are commonly subsumed under the heading of *character simplification* (cf. Hu 2015). With respect to the results of these changes, Zev Handel (2013: 21) comes to the devastating conclusion that “today’s simplified character script cannot be categorized as an effective reform by any reasonable metric – it is only simpler in the crudest of senses”. To provide a bit of background, in 1956 and 1964, in the People’s Republic of China, a two-staged simplification of the writing system of Chinese was undertaken. Its goal was to make the writing system easier to learn and use (especially to memorize). Reducing the number of strokes that characters are composed of was believed to achieve this, which, on a grand scale, was thought to result in higher literacy rates (cf. Handel 2013: 39). Handel reconstructs in detail how the simplification, as an external orthographic intervention, does not suit the graphematic properties of the Chinese writing system. On the contrary: in many cases, the simplification opacified the system’s natural graphematic features and “increased systemic complexity” (Handel 2013: 43). Only some of Handel’s numerous examples will be discussed below to underscore how orthographic regulation can be unnatural when it interferes negatively with the graphematic module.

The following examples pertain to Chinese graphemes of the so-called *xíngshēng* 形聲 group (cf. Gong 2006: 47f.) that are composed of both a semantic and a phonological component (cf. Section 2.2). They constitute the most common type of Chinese graphemes. Take, as an example, traditional <讓> *ràng* ‘let, make’, which was simplified to become <让>. What Handel (2013: 43f.) names *incomplete replacement of phonetic elements* is essentially the failure to analogously simplify other (admittedly, low-frequency) graphemes that include the same ‘phonetic element’ (or phonological component), which in this case is <襄> *xiāng* ‘assist’. The graphemes that are, following the reform, licensed as orthographically correct thus do not exhibit systematic relationships with each other. The ironic thing, here, is that they did before. Aside from several other inconsistencies that affect graphemes’ phonological components, Handel (2013: 45) also mentions positional problems with radical simplification: for example, the grapheme <言> *yán* ‘speech’, as an independent grapheme, was not simplified. In its use as a semantic component, however, it was, and <言> became <讠> in its simplified version. Accordingly, <語> *yǔ* 'language' became <语>, <課> *kè* 'course' became <课>. However, when positioned at the bottom of a grapheme, the semantic component was not simplified: <警> *jǐng* 'warn' stayed the same, as did <譬> *pì* 'example'. Handel (2013: 45) concludes that "what in the traditional system is a single element [...] with consistent semantics has become two distinct elements [...], both of which must be learned, and both of which occur with the same functional role; the choice is determined positionally". Admittedly, as Handel notes with respect to this example, this type of complex positional allography (cf. Section 2.3) already existed in the writing system of Chinese. Thus, these newly introduced allographs "serve to exacerbate existing patterns of inconsistency rather than introducing new types of inconsistency" (Handel 2013: 45). This is obviously not something orthographic regulations should do: they should neither introduce new inconsistencies nor increase existing ones. What orthographies should do instead is use the opportunity of external regulation to curb existing inconsistencies. Simply put, if it is necessary to interfere with a naturally grown writing system at all, then, for this interference to be justifiable, it should function to decrease existing unnaturalness in the graphetic and graphematic modules.

One of the questions Handel raises is whether a writing system with an internal structure such as that of Chinese can even be reasonably simplified in the course of an orthographic reform and still keep its structure and its type, i.e. not be supplanted by a syllabographic or segmental phonographic writing system. He affirms this and proposes several steps that could be taken to achieve this: First, the consistency of the phonological components should be increased. This could be done by agreeing on one phonological component for each of the phonotactically legal phonological syllables of Modern Standard Mandarin, of which there are approximately four hundred (when tone is disregarded). For comparison: right now, there exist about eight hundred phonological components. Since many of the four hundred phonotactically possible well-formed syllables of Mandarin simultaneously serve as the signans of multiple morphemes, homophonous morphemes should be disambiguated by semantic components. The two hundred existing semantic components of the current system should be regularized, and importantly, their positions should be fixed in that they should only occur above or to the left of phonological components.

The rendering of Handel's sophisticated suggestions here is simplistic but suffices to highlight that in sum, his proposed steps would result in a "highly transparent" (Handel 2013: 52) system in the sense of transparency (as discussed in Section 6.4). Additionally, the transparency of the subsegmental components would render the system positionally and compositionally transparent as well. If a reform were to adhere to these suggestions, top-down orthographic regulations would affect the graphematic module (as did the actual simplifications) and, specifically, increase the linguistic fit as well as – quite possibly – the processing fit.[217]

**217** Handel (2013) puts forward his suggestions against the backdrop of modern psycho-

Note that Handel (2013: 54) does not term his proposal 'simplification' but *regularization*. It is worth mentioning that due to naturalness conflicts, other aspects that were taken into account in the actual simplification of Chinese graphemes would need to be neglected in such a regularization: since Handel abandons the practice of stroke reduction altogether, for example, the question of visual complexity and general graphetic naturalness is not touched upon by his suggested regularization. Furthermore, the proposal of a standardized set of four hundred phonological components based on the phonological syllables of Modern Standard Mandarin "would be offset to an unknown degree by a loss of pan-dialectic applicability of the script" (Handel 2013: 56). Thus, such unidialectal phonological components (cf. Section 8.3 for the problem of choosing one variety as the basis for a written standard) would possibly not be broadly accessible to the members of the Chinese writing system who speak a different variety than Mandarin, and the sociocultural fit would decrease. This emphasizes that any intervention must set priorities and decide which competing forces that are in a naturalness conflict are supported and which are simultaneously neglected.

In sum, what this example shows is that external orthographic regulation can decrease the graphematic systematic fit by 'destroying' or opacifying systematic features at the expense of other priorities (such as, in this case, the graphetic processing fit).

linguistic research and underlines that his proposed changes highlight the features of the Chinese writing system that have been found to facilitate processing.

# 6 Linguistic fit

Writing systems are semiotic systems. Specifically, their core, the graphematic module, semiotically relates units of language with visual units. The human interpreter, acting either as a reader or a writer, is actively involved in the semiosis that conceives of graphemes as signs. Accordingly, the semiotic relations that constitute the graphematic module are not only part of the writing system but also a crucial part of its users' knowledge. Notably, it is not only segmental graphemes but rather units of all graphematic levels – including subsegmental graphematic levels and larger levels such as the level of the graphematic word – that can be interpreted as signs. This semiotic nature of writing systems is reiterated here because it makes possible an evaluation of the "quality" of their inherent semiotic structures. This is treated under the heading of the so-called *linguistic fit* (cf. also Sariti 1967: 29). In essence, it aims to provide an answer to the question: how well does a given writing system fit the language it is based on? The naturalness parameters of Natural Morphology that are concerned either with the relationship between the signans and signatum of a sign or with the properties of the sign as a whole can be transferred to and reconceptualized for a graphematic context (cf. Section 4.2.3 for preliminary examples), which constitutes the heart of the present chapter.

The linguistic fit is solely descriptive. In other words, while it can be used to assess the semiotic structures of writing systems, it is not informative with respect to how these structures actually affect cognition or, more generally, how users process writing systems. This is evaluated in the next step, in the context of the processing fit (cf. the next chapter). Notably, in original Naturalness Theory, these two fits appear to be inherently merged: in Natural Morphology, it was explicitly assumed that the semiotic structure of signs affects cognitive processing. Thus, structure and processing were claimed to be inseparably and directly linked. A "good" semiotic structure was believed to automatically equal a good cognitive fit that facilitates processing. In the present proposal of a functional theory of writing, this correlation of the linguistic and processing fits is no longer regarded an axiom. Thus, it remains to be shown whether (and how) semiotic structure – and particularly, which of its facets – actually bear on the processing of written signs.

A broader and more neutral term for the linguistic fit would be *semiotic fit* given that it is interested in the graphematic module's semiotic structures. The choice to call it *linguistic fit* foregrounds the specific question of whether a writing system fits a given language. In theory, the opposite question – i.e. "does a language fit a writing system?" – could also be asked. However, since language is not only ontogenetically and phylogenetically primary but writing as a modality is also fundamentally dependent on language (and writing systems are actually created/adapted with a specific language in mind), the perspective of language is taken – hence *linguistic fit*.

The investigation of linguistic naturalness at the graphematic level is located in part at a different level than Natural Morphology, the subbranch of Naturalness Theory it is (primarily) modeled after. As a subsystem of language, morphology is a primary sign system, which makes morphemes primary signs. As such, they relate a phonological representation (= signans) with a meaning (= signatum). By contrast, morphographic writing systems such as Chinese are secondary sign systems: in a grapheme such as Chinese <家> *jiā* 'house', the basic shape corresponds with a morpheme, which is itself a sign. Thus, the grapheme is a sign of a sign. In phonographic systems such as German, too, a sequence of graphemes such as <Haus> 'house' relates to a morpheme, rendering the graphematic word <Haus> also a sign of a sign. However, at the segmental level, the individual graphemes that constitute <Haus> have a different status than Chinese graphemes: they relate basic shapes to phonemes, which, unlike morphemes, are not signs (cf. Section 4.2.3). Accordingly, much like the morphemes studied by Natural Morphology, phonographic graphemes are primary signs. Several naturalness parameters such as *pictography* cannot (easily) be evaluated for phonographic graphemes: for example, the phoneme /s/ that <s> corresponds with is not a sign since it lacks a signatum, i.e. it has no 'meaning' that could be iconically depicted, which in turn means that it cannot be assessed whether <s> is iconic. Furthermore, as a mere signans, /s/ (or its materiality) can also not be iconically depicted itself (with exceptions as in Korean Hangul, cf. Section 6.2). The situation is expectedly different for the above-mentioned morphographic grapheme <家>. Its signatum is a morpheme that has a meaning. Therefore, whether this grapheme is iconic can be assessed by evaluating whether the basic shape's visual form resembles the (prototypical) shape of the object that the morpheme's meaning refers to, i.e. a house.

The fundamental difference between 'meaningless' phonography and 'meaningful' morphography boils down to the fact that for phonographic writing systems, several naturalness parameters cannot be evaluated, at least not for individual graphemes. Notably, with respect to larger graphematic units such as the graphematic word <Haus> 'house' in which the string of graphemes does relate to a meaning-bearing linguistic unit, such evaluations are possible also in phonographic systems. By contrast, in morphographic systems such as Chinese, the semiotic naturalness parameters of Natural Morphology can be applied straightforwardly. This is not particularly surprising, since the theoretical and methodological apparatus of Natural Morphology was tailored for morphology, and morphographic systems are based on morphology.

In the following subsections, I will sketch how the linguistic fit of writing systems can be evaluated one parameter at a time by discussing examples from diverse writing systems, predominantly Arabic, Chinese, German, Korean, and Thai, but also others. The parameters will be characterized concisely, and the focus will be on the question of which configurations on these parameters are more natural than others. In this context, it is necessary to separate the different levels at which certain parameters can be natural, i.e. universally, type-specifically, or even system-specifically. In doing this, this chapter also offers tools that help in systematically investigating the question of whether "every language gets the writing system it deserves" (Frost 2012: 266); however, it does not provide a definite answer to said question.

Several different categories of parameters need to be distinguished: first, *paradigmatic parameters* vs. *syntagmatic parameters*. Transparency, for example, compares graphemes *paradigmatically* within a system by investigating how transparently they relate to linguistic units such as phonemes, syllables, or morphemes. By contrast, it is incapable of capturing phenomena that occur only in larger graphematic sequences, an example of which is positional incongruency, i.e. when a graphematic sequence <abc> relates to a phonological string whose constituents exhibit a different order such as /bac/. Here, the individual graphemes are (paradigmatically) transparent but there is unnaturalness on the parameter of positional transparency, which can be captured only *syntagmatically*.

A second necessary distinction is *intersystemic parameters* vs. *intrasystemic parameters*. Depending on the perspective taken, many of the parameters listed in the following can be both. For example, when the paradigmatic transparency of the German graphemes <f> and <v> is compared, transparency is interpreted intrasystemically. However, when the transparency of German graphematics is compared with the transparency of Thai graphematics, transparency is conceived of intersystemically. Some parameters can be reasonably interpreted only in one of those ways: indexicality, for example, evaluates how natural the indexical relations are between certain elements of texts, e.g. how spatial proximity or distance transparently signal or opacify semantic textual coherence. This parameter can be applied most fruitfully in a comparison of texts, specifically textual arrangements or layouts. Arguably, at this higher-order graphematic level, different writing systems share (almost) the same resources to arrange texts, which means they are equal from an intersystemic perspective. Intrasystematically, however, different texts from one system – e.g. German – can be compared (or, at an even lower level, different instances of indexicality within one text can be compared intratextually).

In sum, this chapter will showcase parameters that allow both a description and a comparison of writing systems and thus sets the stage for the investigation of how these parameters affect the processing of writing systems.

## 6.1 Unit of representation

The first parameter to be discussed here evaluates which type of linguistic unit a writing system's graphemes should primarily relate to. In theory, all closed inventories are candidates, including phonological features, phonemes, syllables, and, to some degree, morphemes. By contrast, words, sentences, texts, or discourses are not suited to be the so-called unit of representation (cf. Meletis accepted b).

At this point, it is tempting to resort to alphabetocentric arguments, which claim that a writing system's ability to relate to segmental linguistic units renders it more efficient. Likewise, one could advocate the syllable as the most natural unit of representation given that evidence from processing suggests it is the unit that is most efficiently processed (cf. Section 7.2.1). However, on the one hand, graphematic modules can also efficiently represent their underlying language systems without necessarily offering segmental (i.e. phonemic) information, and on the other, non-syllabographic writing systems are, depending on their specific makeup, potentially also processed efficiently. Thus, from a purely descriptive point of view, no such thing as a unit of representation that would be natural universally exists. Instead, the crux of the question of whether languages 'get the writing systems they deserve' (cf. Frost 2012: 266) lies not at the universal but the typological and system-dependent levels of naturalness. Accordingly, it is the type a language belongs to and, to varying degrees, also its idiosyncratic features that determine which linguistic unit graphemes should correspond with in order to maximize the linguistic fit of a writing system.

The claim that language typology and writing system typology interact in crucial ways should be elucidated in more detail. Firstly, one can observe that several features of languages can influence which unit of representation is most natural for a given language. Some of them are closely related to classifications made by morphological and phonological typology. They include (but are not limited to) the size of a language's phoneme inventory, the number of phonotactically licensed syllables, the degree of invariance of morphemes' shapes (which involves morphonology), the existence of consonant clusters and the associated degree of syllable complexity, the length of words, homophony, infixation and introflection, cumulation, fusion, and the existence of tones. Many of these features interact with each other or are even determined by one another; take, for example, the number of possible syllables and homophony (see below). I want to highlight the general relevance of such typologically relevant features by discussing two specific writing systems: Chinese and German.

Chinese is most often classified as an isolating language (cf. Whaley 1997: 129). It is uncontroversially an analytic language, typical features of which are (1) predominantly monosyllabic morphemes, (2) lexical tone, (3) extensive use of function words, and a (4) relatively fixed word order. All these features are found in Chinese. Additionally, the Chinese phoneme inventory is relatively small and its phonotactics exhibit a low degree of complexity. This yields a relatively small number of possible well-formed syllables. And given that Chinese morphemes are (largely) monosyllabic, there is a large degree of homophony at the level of morphemes. Additionally, the phonological representation of morphemes is almost maximally constant, mostly because of the lack of inflectional morphology. The large degree of homophony is disambiguated by the existence of four lexical tones that roughly multiply the number of possible syllables by four, and it is further disambiguated by the fact that most words are bisyllabic. Now, the question of whether Chinese could be written alphabetically rather than morphographically has already been mentioned briefly (cf. Section 4.2.1). Aside from claims to the alphabet's superiority and other external (sociocultural) factors, this question can (and arguably should) be settled on linguistic grounds alone – which is precisely what the *linguistic fit* does. Writing Chinese alphabetically would reduce the number of graphemes drastically, as the unit of representation would be the phoneme instead of the morpheme, and there are relatively few phonemes in languages when compared with the number of morphemes. Notably, a phonographic writing system for Chinese would need to solve the problem of homophony. Indeed, in an alphabet that is phonographically transparent, i.e. lacks any morphography, homography would be pervasive, since in a completely transparent alphabet, everything that exhibits the same phonological representation is also written in the same way. To disambiguate between different morphemes, i.e. different meanings, secondary morphographic information could be introduced to such an alphabet, although how this could be reasonably achieved is questionable since a single Chinese syllable often serves as the phonological representation of ten or more morphemes.

A hypothetical syllabographic writing system for Chinese encounters many of the same problems. The grapheme inventory would also be smaller given that there are fewer syllables in Chinese than morphemes. In both alphabetic and syllabographic renditions of Chinese, the inclusion of tone – via tone markers, for example – would increase compositional transparency (cf. Section 6.6) but would still not disambiguate the meaning of morphemes that are homophonous even when tone is accounted for. The bottom line is that a phonographic writing system for Chinese would introduce a massive degree of morphographic opacity that would have to be dealt with by introducing secondary morphography. This seems counterproductive: why switch from a morphographic to a phonographic system to then only make it more morphographic again?

At first glance, the sole apparent benefit of a phonographic system would be, from a descriptive point of view, a smaller number of units, i.e. descriptive economy. Crucially, for processing, this equals fewer units that must be acquired by users to be literate in the system. By contrast, the morphographic system of Chinese boasts a huge grapheme inventory, which is, however, morphographically largely transparent. Interestingly, it includes also a certain degree of phonographic transparency via the phonological components incorporated into its morphographic graphemes. One of the features of Chinese that could potentially be better accommodated in a phonographic system is lexical tone, which is not represented in the morphographic system of Chinese. Based on these considerations, it is a justified question whether the cognitive cost associated with a large grapheme inventory is greater than a large degree of homography and morphographic opacity. It is, however, not a question that is asked in the context of the descriptive linguistic fit.

German, by comparison, is a synthetic, more specifically a fusional language. It boasts many consonant clusters both in initial and final syllable positions, and its syllable structure, in general, is complex. This makes a syllabographic writing system unfeasible, as it would make necessary an enormous number of graphemes. Due to morphological processes such as ablaut as well as phonological processes such as final devoicing, morphemes can change their form to some degree. Additionally, grammatical morphemes are cumulative and morphemes such as {-st}, for example, encompass multiple functions, in this case 2nd person and singular. And these are just a few of the features that would make a morphographic writing system for German cumbersome. In sum, it appears that the alphabet is indeed the best linguistic fit for German.

Admittedly, these were merely two brief thought experiments. Nevertheless, they effectively underline how a language's type as well as the specific features of the language in question are highly relevant in evaluating which unit of representation is most natural. What must be emphasized, of course, is that the unit that a writing system's graphemes actually relate to is not always determined on linguistic grounds, since factors such as script adoption (cf. the introduction to Part II), specifically the type of the writing system the adopted script was originally used for, as well as sociocultural factors often come to the forefront in the context of this decision.

## 6.2 Iconicity

Ahead of an analysis of graphematic iconicity (cf. also Glück 2011; Stetter 2011), some of the basic semiotic facts about iconicity shall be presented; they are based on Nöth (2001), who himself treats Peirce's views on iconicity.

As one type of sign, the icon is "defined according to the relationship between the sign and its object" (Nöth 2001: 18). More specifically, a sign is regarded as an icon if it resembles the concept of the extralinguistic referent it stands for.[218] Thus, prototypical examples of icons are photos, portraits, and realistic paintings. Icons are one of three types of signs, the other two being indexes (for indexicality, see below) and symbols. Concerning these three types, it is important to note that "[e]very language sign [including signs of writing, D.M.], even an iconic or an indexical word is a symbol" (Nöth 2001: 19), with symbols being characterized by an arbitrary relationship between the sign and its object. This underlines

**218** As Nöth (2001: 20) argues, for Peirce, an object "is no object of external reality, no object that exists independently of the sign". The object is rather something "merely imaginable". Much like the *interpretant*, i.e. the sign constituted by the semiosis that is carried out by the interpreter, the object is also a sign, but it "precedes our interpretation [and] is less developed than the interpretant of the sign". For our purposes, it suffices to treat the object as an abstract concept of extralinguistic referents. Thus, the morpheme {tree} does not refer to any one tree that exists in reality or even the sum of all trees that exist in reality but to the mental concept humans have of a tree. Notably, this mental concept is, of course, still intricately linked to and affected by our perception of trees that exist in reality.

that these sign types are not mutually exclusive; more specifically, it means that all signs are symbols. Consequently, every evaluation of linguistic iconicity equals an evaluation of iconicity in symbols.

Notably, linguistic icons are not genuine icons but so-called *hypoicons*. A hypoicon is "only similar to its object, and it shares only some of its features with its object" (Nöth 2001: 19). Three subtypes of hypoicons are distinguished: (1) *images* exhibit a similarity to their objects; an example is onomatopoeia as evidenced by words such as "cuckoo"; (2) *diagrams* bear a structural or relational similarity to their objects, an example of which is *ordo naturalis*, the 'natural order' in "Veni, vidi, vici", where the ordering of linguistic signs in the utterance reflects the temporal sequence of how the described events unfolded in reality; lastly, (3) *metaphors* are least iconic, since in them, "[t]he ideas conveyed by the sign and the idea of its object are mediated by a third idea, the *tertium comparationis* between the tenor and the vehicle of the metaphor" (Nöth 2001: 21, emphasis in original). Specifically, the *vehicle* (or *source*) lends some of its attributes to the *tenor* (or *target*). For example, in the well-known Shakespearean metaphor "All the world's a stage", 'stage' is the vehicle whose attributes are transferred to 'world', the tenor.

Another basic distinction that proves crucial for an evaluation of graphematic iconicity is that between (A) *exophoric* and (B) *endophoric* iconicity, or, as Nöth (2001) paraphrases them, "form miming meaning" and "form miming form", respectively. The examples provided so far are instances of (A) exophoric iconicity. Here, signs "serve as icons of the world, and iconicity consists of a signans-signatum relationship" (Nöth 2001: 21f.). As this type of iconicity is constituted by signs *referring* to something that is external to them (i.e. something that exists outside of language), it is sometimes also referred to as *referential iconicity*. The second type, (B) endophoric iconicity (also known as *relational iconicity*), by contrast, has to do with "relations of reference *within* language" (Nöth 2001: 22, my emphasis). Two further subtypes of endophoric iconicity are distinguished based on whether they function (B1) *syntagmatically* or (B2) *paradigmatically*. (B1) Syntagmatic endophoric iconicity is relevant for intratextual analyses: for example, repetitions, parallelisms, alliterations, and rhymes are in syntagmatic relationships within texts and can be syntagmatically iconic. (B2) Paradigmatically – and more importantly for a functional theory of writing –, endophoric iconicity is concerned with intrasystemic iconicity, i.e. iconicity within an entire system (rather than just a text based on that system). As Nöth (2001: 23) notes, endophoric iconicity is closely related to the concept of diagrammaticity (see below). To give a grapholinguistic example: That the graphemes <m> and <n> are materialized by the visually similar basic shapes |m| and |n| and simultaneously relate to phonologically similar units (nasals that only differ in one phonological feature) in virtually all alphabets in which they occur is an instance of endophoric diagrammaticity.

Based on these preliminary considerations, three types of iconicity are central for an investigation of graphematic naturalness: (a) (exophoric) imagic iconicity in the form of *pictography*, (b) exophoric diagrammaticity in the form of structural relations between a grapheme and the object (or rather the linguistic unit) it relates to, and (c) endophoric diagrammaticity as in the above-mentioned example of <m> and <n>, where a diagrammatic (and indexical, see below) relation is established between two units of one system. Examples of all three types as well as some examples of more marginal types of iconicity not mentioned here will be discussed in the following.

(a) In a grapheme (or larger graphematic unit, for that matter) that is exophorically imagic, the signans (the basic shape) visually resembles the object that the signatum (the linguistic unit) refers to. This is the case especially in some morphographic graphemes whose signantia visually resemble the signata's referents. In the case of writing (or other types of visual communication), this imagically iconic relation is commonly referred to as *pictography*.[219] In the Chinese grapheme <木> relating to the (Chinese) morpheme {TREE}, which, in turn, has the concept of 'tree' as its signatum, the basic shape |木| to some degree visually resembles a tree. From a diachronic perspective, pictography has decreased so drastically that synchronically, it is a rare feature of writing systems. This is reflected by the central graphetic features of arbitrariness and abstractness (cf. Section 7.1.1). Pictography was, thus, a much more prevalent feature in ancient writing systems than it is in modern writing systems. Take Egyptian hieroglyphs or Maya script as examples, both of which retained pictographic components but, similar to Chinese, introduced arbitrary phonographic components as well (cf. Kim 2011: 187). The historical decrease of iconicity is caused partially by human processing pressure, specifically the fact that abstract shapes are easier to produce (cf. Section 7.1). Boltz (1986: 426) mentions another reason for the decline of pictography, which is dependent on the structure of language and is thus relevant for the linguistic fit of pictography:

> Clearly such a primitive device such as drawing a picture to represent the intended word will soon prove unable to cope with even the simplest abstractions, much less with the manifold semantic complexities of the whole language. At this initial stage if something could not be depicted directly, it could not be written.

Accordingly, pictography is rather unnatural linguistically since, as Boltz observes, it cannot handle the semantic richness of language (cf. also Tversky 1995: 34). How should abstract notions such as "freedom" or "love", not to mention function words (or complex morphology, for that matter), be depicted pictographically?

Another special characteristic of pictography is that it lies in the eye of the beholder as it is constituted by the interpreter (cf. Ding 2005: 277). As such, it is not a truly objective descriptive category (if there even is such a thing)[220] since it always depends on the interpretation of humans. Consequently, pictogra-

**219** Cf. Boltz (2006) and Jespersen & Reintges (2008) for criticisms of the term and concept *pictography*. Like Behr (2010: 291), I use the term here "for easier cross-reference to the sinological tradition" and other philological traditions as well as to make pictography visible and comparable across writing systems.

**220** One could argue that description is *always* subjective as it is always, in some way, determined by the descriptor.

phy is relevant to both the linguistic and processing fits. But while it is constituted by users' perception, this alone does not reveal anything about how it affects processing. In this vein, Xiao & Treiman (2012) tested in their study which characters users regard as pictographic in the contemporary writing system of Chinese. Participants who did not speak or read Chinese were first presented a word or phrase in English and were then given two distinct Chinese basic shapes (made up of an equal amount of strokes). Participants then had to decide which of the shapes corresponded to the meaning of the word given in English. Notably, as the participants neither spoke nor read Chinese, they were incapable of linking the shapes to the respective Chinese morphemes they relate to graphematically. However, prior to being instructed to choose a fitting shape, they were given the information "that Chinese characters tend to look like what they mean and that they could look for similarity between the appearance of characters and the meaning" given in English (Xiao & Treiman 2012: 957).

The study's results suggest that only 15 of the over 200 basic shapes that were presented are deemed pictographic at above chance level. These are |爪| for 'claw', |凹| for 'concave', |内| for 'inner-', |贝| for 'shell', |气| for 'air', |川| for 'river', |个| for 'individual', |丁| for 'man', |小| for 'small', |耳| for 'ear', |田| for 'field', |三| for 'three', |门| for 'door', |雨| for 'rain', and |云| for 'cloud' (cf. Xiao & Treiman 2012: 957). One of them, |三|, is actually an example of a different type of iconicity, exophoric diagrammaticity (see (b) below); it is the only one associated with an abstract concept.

For phonographic graphemes, the situation is more complicated. Here, as outlined above, phonemes or syllables serve as the graphemes' signata. Since phonemes or syllables (if they are not already morphemes) do not have a referent (or an object they refer to), phonographic graphemes would have to visually resemble the phoneme itself in order to be pictographic. Arguably, there is no way of truly iconically representing an acoustic datum (or a theoretical abstraction of said datum) in a visual modality as visual vs. acoustic data transmission differ fundamentally. Diagrammatically (see below), such a representation would be possible by means of spectrograms or oscillograms that visualize a phoneme's (more specifically phone's) or syllable's features in the context of the continuous stream of speech but not the phoneme (phone) or syllable itself.

Concerning phonographic pictography, the script used for the writing system of Korean, Hangul, is a special case. Its basic shapes are formed to pictographically resemble the place of articulation of the phonemes they relate to (cf. Kim 2011). It is this property of the Korean writing system that serves as the basis of its typological classification as *featural* (cf. Sampson 2015; Section 2.7): the smallest type of linguistic information that is represented graphematically is the shape of the speech organ and the articulatory gesture during a phoneme's production (cf. Kim 2011: 181). This, in terms of phonological features, is the place of articulation. Even if the pictography exhibited by Hangul does not play a major role in its use, i.e. the processing of the Korean writing system – to my knowledge, there are no studies that show whether L1 or L2 users of Korean are aware of the iconicity while reading, and it is questionable whether this feature affects the processing fit unconsciously –, this pictography also leads to a high degree of endophoric diagrammaticity in the writing system (see below). This is because the graphemes' exophoric iconicity automatically results in a situation in which shapes that refer to similar phonemes (e.g. bilabials) are visually similar.[221] In other words, linguistic similarity is reflected by visual similarity.

(b) The second type of iconicity relevant in an analysis of writing is exophoric diagrammaticity. Obvious examples also come from Chinese: notably, exophorically diagrammatic graphemes, classified as a group called *zhǐshì* 指事 'indicating things' according to the six-part *liushu* nomenclature (cf. Gong 2006: 40), are relatively few. The most-cited examples include the graphemes <一>, <二> and <三> that represent the numerical concepts 'one', 'two', and 'three', respectively, as well as <上> and <下> that refer to 'above' and 'below', respectively. These concepts of different quantities or different locations/directions cannot be represented in a straightforward pictographic manner since they do not refer to concrete objects the same way 'tree', 'door', or 'claw' do. However, diagrammaticity is still possible: the increase of numerical value can be represented visually by the increase of the number of strokes, and spatial concepts can be visually represented by means of a structural analogy (the full-length horizontal stroke being below, which represents 'above', and the full-length horizontal stroke being on top, signifying 'below'). This type of diagrammaticity can also be combined with the pictography described in (a), for instance in the grapheme <本> *běn* 'root', where a stroke at the bottom of the basic shape that pictographically represents 'tree' indicates diagrammatically something at the bottom or below a tree: a root.

(c) The third type of iconicity relevant to the linguistic fit is paradigmatic endophoric diagrammaticity, also referred to as relational diagrammaticity. It is arguably the most important type of iconicity for the synchronic analysis of modern writing systems. As it subsumes a range of heterogeneous phenomena, it is a very broad type. In general, it refers to the fact that the structural difference between the signata of two (or more) signs is reflected by a structural difference in their signantia. When singular *artist* is pluralized as *artists*, for example, it is not only the conceptual meaning that "increases" (from 'one artist' to 'more than one artist'), but also the signans, since the grapheme <s> is added to the graphematic word and it thus increases in length. In English, this corresponds with the increase of phonological material in the spoken form of the plural; this renders the written

**221** Importantly, these two phenomena are not mutually dependent on one another: Korean graphemes could, in theory, also be diagrammatic if the individual shapes did not pictographically represent the place of articulation but were instead designed arbitrarily. The decisive criterion for diagrammaticity is that similar linguistic units are depicted similarly in visual terms, and this is of course possible also for arbitrary basic shapes and their relations. In the case of Korean, however, it is *because* of the pictography that diagrammaticity is constituted, i.e. diagrammaticity is a byproduct of pictography. See the diagrammaticity of Roman script below for an example that is independent of pictography.

form of the plural less evidently a form of true graphematic diagrammaticity, as it could be interpreted as being merely a reflection of the diagrammaticity already evident in phonology. Take French, however, where *artiste* is pluralized as *artistes*, both of which are pronounced /aʁ.tist/. In this case, thus, the grapheme <s> serves as a plural marker and establishes a diagrammatic relationship dependent on the singular and plural signs as two elements of the graphematic subsystem of French. This relationship is independent of the respective phonological representations, which are not diagrammatic. Such examples of diagrammaticity caused by pluralization concern – in non-morphographic writing systems – larger graphematic units (such as graphematic words), but diagrammaticity exists also at lower levels between individual graphemes of a writing system. This type of segmental graphematic diagrammaticity was introduced above with the example of Korean Hangul. In the following, I want to focus on it in more detail by discussing the groundbreaking work of Beatrice Primus.

In her "featural analysis of the Roman alphabet", which she also refers to as "grammar of letters", Primus (2004; 2006) analyzes the elementary forms of the lowercase basic shapes of Roman script first graphetically and then graphematically, focusing on the linguistic, and more specifically, phonological values of visual features (for a critical overview of Primus' approach, cf. Meletis 2015: 66–76). Specifically, Primus aims to show how certain visual features of elementary forms that make up basic shapes correlate with phonological features and how the sum of the visual features' phonological values equals the phonological feature makeup of the phoneme that a given basic shape relates to.

The first step in Primus' analysis is the segmentation of lowercase basic shapes into their respective elementary forms (cf. also Section 5.1). The elementary forms resulting from this segmentation are not all of equivalent status. Following Brekle, Primus identifies two different types of elementary forms: *heads* (in Brekle's terms, *hastas*) and *codas*. In |d|, the vertical stroke |l| is the head and the smaller curved stroke |c| is the coda. One of Primus' central observations is that heads are always vertical and obligatory and can – but are not required to – exhibit the feature [+length]. By contrast, codas are optional, can be – but are not necessarily – vertical, and never exhibit the feature [+length] (cf. Primus & Wagner 2013: 41, see also Figure 31). At this point, the specific form-function correlations as described in detail by Primus shall not be discussed, but I want to argue, in broad strokes, how impactful her discovery is.

When it comes to the phonological feature of *place of articulation*, for example, the features of basic shapes' heads are revealing: basic shapes with a long and straight head (as in |p, t, k, b, d|) are in graphematic relationships with plosives, basic shapes with long but curved heads such as |f, v, w, s| with fricatives, and basic shapes with non-long heads as in |m, n, r| with sonorants (cf. Primus & Wagner 2013: 44). Furthermore, in the ensuing graphemes, at least one feature of their basic shapes' heads marks their diagrammaticity: [+length] (for a discussion of the role this feature plays for the demarcation of graphematic syllable structure in alphabets, cf. Section 2.4).

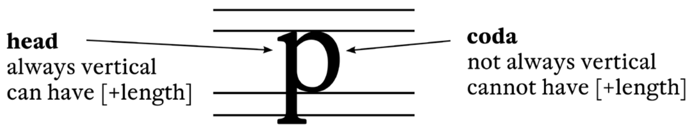



FIGURE 31. Head and coda structure of |p|, from Berg, Primus & Wagner (2016: 339)

One of the criticisms that have been voiced against Primus' analysis is based on the fact that Roman script is used for many different alphabets; for this reason, her assumed form-function correlations, Rezec (2010: 345f.) argues, cannot possibly account for all of the graphematic relations that the basic shapes partake in across alphabets. Concerning this point, it must be countered that the sum of the phonological values that a basic shape's elementary forms add up to does not equal the full phonological representation of the phoneme the basic shape relates to (in a specific writing system such as German) but instead only an *underspecified* phonological representation. In other words, a basic shape's featural configuration corresponds with the "canonical phonological value", i.e. an underspecified phonological feature bundle. This, ultimately, allows for basic shapes to be in graphematic relationships with several phonologically related phonemes within and across alphabets (cf. Primus 2006: 21). Primus & Wagner (2013: 46) illustrate this with the example of |j|, whose head is long, left-oriented, and not straight. Based on the form-function correlations assumed by Primus, these features merely point to a fricative produced in the back of the oral cavity. The exact phonological value of the basic shape, however, can vary across writing systems: in German *jeder* 'everyone' it corresponds with /j/, in French *jean* 'jeans' with /ʒ/. Notably, there might be exceptions to this: in the – admittedly unlikely – event that Roman script is adopted for a writing system *without* any of the graphematic relations from its donor writing system (cf. the introduction to Part II for pure script adoption), graphematic relations could be (re-)assigned randomly and the diagrammaticity constituted by form-function correlations outlined above would be 'destroyed'.

What this example underlines is that the systematic fit of Roman script (cf. Section 5.1), which was evaluated entirely independently of the graphematic relations that the basic shapes are parts of in various alphabets, can be (and is) diagrammatic when it is functionalized graphematically: this is the case when the visual similarity of basic shapes (stemming from the fact that they share the features of a given script as a visual system) is translated to graphematic similarity. In a nutshell, thus, diagrammaticity is established when visual similarity indicates linguistic similarity. If, by contrast, the systematic fit of a script is not or only to a small degree linguistically functionalized, the possibilities that existing graphetic naturalness offers are 'wasted' at the graphematic level. Consider Thai: many basic shapes in Thai exhibit a striking visual similarity and many pairs or groups of basic shapes within the script differ only minimally (cf. Cooper 1996). In some cases, visual similarity is functionalized, for example in the graphemes <บ> and <ป> that relate to /b/ and /p/, respectively, i.e. phonemes that differ in one feature, the feature [±voiced]. By comparison, the exact same visual difference between two basic shapes is not functionalized in <พ> and <ฟ> that relate to /p$^h$/ and /f/, respectively. Thus, in these two pairs of graphemes, an analogous visual modification does not reflect an analogous linguistic relationship between the phonemes that the graphemes relate to. Thus, this specific visual modification is not diagrammatic. Note, however, that the latter two graphemes <พ> and <ฟ>, which are low-class consonant graphemes, are in a diagrammatic relationship with <ผ> and <ฝ>, which also represent /p$^h$/ and /f/, respectively, but are high-class consonant graphemes.[222] When considering only these four graphemes, 'orientation of the loop' is diagrammatic, with [loop on the right][223] signaling *high class* and [loop on the left] signaling *low class*. Both diagrammaticity and non-diagrammaticity of visual differences abound in the writing system of Thai; notably, most instances of diagrammaticity are only local, such as the latter example provided here, meaning they pertain merely to pairs or groups of graphemes but not the entire writing system.

Diagrammaticity is intricately linked to another naturalness parameter, indexicality (see below). In the German writing system, for example, graphemes that are embodied by basic shapes with straight strokes as descenders or ascenders (such as <p> and <k>) index and evoke each other.

In conclusion, it must be emphasized that the above-mentioned phenomena subsumed under the heading of iconicity cannot always be separated neatly and actually often occur simultaneously. Take the Chinese basic shapes |木| and |森|. Here, there is a (subjectively perceived) pictographic similarity of the respective basic shapes with a single tree and three trees, respectively, which renders these basic shapes pictographic. Simultaneously, they are exophorically diagrammatic since their relationship roughly depicts the increase of real substance (from a single tree to a concept such as 'woods') as it would look like in reality. Finally, at the graphematic level, used as graphemes, <木> and <森> are endophorically diagrammatic because the increase in visual substance from the first to the second signans diagrammatically reflects the 'increase' from linguistic singularity to plurality. While for the linguistic fit of a writing system, exophoric imagic iconicity and exophoric diagrammatic iconicity are only of marginal relevance, endophoric diagrammatic iconicity is central.

---

**222** In the writing system of Thai, there are three so-called *consonant classes* that consonant graphemes are assigned to: high, middle, and low consonants. The class of a consonant grapheme is important for the assignment of tone to the written syllables of Thai (cf. Haas 1956: 10f.).

**223** In typographic terms, the loop in |พ| is called *front-first loop*, whereas the loop in |ผ| is referred to as *back-first loop* (cf. Punsongserm, Sunaga & Ihara 2017: 7).

## 6.3 Indexicality

To describe how the parameter of indexicality can be reconceptualized for graphematics, it is useful to (re)consider how it was defined in Natural Morphology and especially the lesser-known branch of Natural Textlinguistics. In Natural Morphology, so-called "[i]ndexical force is the function of the major/minor proximity relating index and indexed sign both in the phonological sequence and in the hierarchy of content" (Crocco Galèas 1998: 77). Lexemes are seen as indexed signata, and there exist different types of indexing signantia, the most natural of which are derivational affixes since they are contiguous to the base and "denote the lexical value" of the lexemes (Crocco Galèas 1998: 77). Inflectional affixes, for example, are less natural as they do not affect the lexical meaning of lexemes but only their grammatical meaning. Aside from the factor of how much indexes or indexing signantia affect the lexical meaning of indexed signata, the distance to an indexed signatum and the position within a sign play crucial roles for the different naturalness values of indexicality. The central factor for the evaluation of the naturalness of indexicality is heavily influenced by gestalt theory: *spatial contiguity*.

In Natural Textlinguistics, indexicality was identified as one of the most prominent parameters (cf. Fludernik 1996: 322), focusing on intratextual relations between the elements of a text. Here, too, the type of index and its relative location play a role. For example, anaphoric indexes such as personal pronouns that refer back to earlier elements of the text such as proper names (which should not be positioned 'too far' away) are claimed to be most natural. How could, based on its interpretation in Natural Morphology and Natural Textlinguistics, indexicality be modeled for the graphematic module?

Indexicality is, together with aspects of figure—ground (see below), most relevant for the analysis of higher hierarchical levels of writing, i.e. not the level of graphemes but the levels of graphematic words, graphematic sentences, etc. Indeed, evaluating indexicality works best for an intratextual analysis of entire texts. This type of high-level graphematic indexicality was already foreshadowed by Dressler (1989: 48–49) in his treatment of Natural Textlinguistics when he discussed footnotes, endnotes, lists of references, etc., as examples. The position of these elements within the entire text plays a crucial role, and more specifically, their spatial relationship (measured in proximity or distance) with the indexes referencing them in the running text.

To provide an example, a trivial assumption that can be made with respect to what is natural indexically is that footnotes, as indexes or indexing signantia that are referred to in the main text (and simultaneously refer back to the main text), are more natural than endnotes. The reason for this is that to read endnotes, readers must turn the pages to the end of a book, chapter, or article to read the notes, whereas, for footnotes, they "only" have to move their eyes to the bottom of the page. Thus, in the vein of Naturalness Theory, there is less effort involved in processing footnotes than endnotes because they are located *closer* to their indexed sign. This is, of course, just one example, and in fact, it does not concern the linguistic fit of indexicality, which is evaluated descriptively, but instead the processing fit (cf. Section 7.2.3).

A full-fledged theory of textual indexicality must identify all relevant elements within a text as a closed system and, in the next step, investigate their indexical relationships. With respect to distance, as the example of footnotes already implies, proximity is more natural, and with respect to position, anaphoric indexes, e.g. indexes that spatially (and in terms of reading, chronologically) follow their indexed signata, are more natural than cataphoric ones. Additionally, precise indexes are more natural than imprecise indexes. An example of an imprecise index is “see above” while a precise index would be “cf. Chapter 2”. Even more precise are footnote numbers that are positioned directly after the word or sentence that the footnote’s content refers to. All these indexes are, as mentioned, intra-textual, rendering them endophoric. An example of exophoric indexes are references to other works that exist outside of the text, e.g. references such as “Dressler (1989)” in the present text. Again, a specific reference is more natural than something along the lines of “in the literature on Natural Textlinguistics”. Position, distance, and specificity are thus relevant subparameters for the evaluation of the descriptive naturalness of endophoric indexicality.

In a nutshell, indexicality can best be evaluated for higher levels of graphematic organization, mainly the textual level. In this respect, indexicality is rather unique among graphematic naturalness parameters since it serves an intra-textual evaluation of naturalness rather than an evaluation of naturalness in a given writing system (intra-systemic) or across different systems and their resources (inter-systemic). This means that with respect to indexicality, writing systems are equal and it is rather specific texts realized in different writing systems – or texts realized in one writing system – that are not equal because their respective structural makeup can be more or less natural indexically.

## 6.4 Transparency

Transparency is the semiotic parameter that is most prominent in the psycholinguistic literature on writing (specifically on reading processes) focusing on processing (which will be discussed in the context of the processing fit, cf. Section 7.2.4). It deals with the analytical direction *signans* → *signatum*, which, in the case of graphematics, means visual unit (basic shape) → linguistic unit (phoneme, syllable, morpheme, …). The central question is: does a given basic shape graphematically relate to one or more linguistic unit(s) in a writing system? In the writing system of German, for instance, the grapheme <f> is transparent. It is always recoded as /f/ – independently of the graphematic context it occurs in. This makes it maximally transparent or, in other words, /f/ exhibits the most natural value on the parameter of transparency.

In their study, Neef & Balestra (2011) develop a measure of quantifying graphematic transparency (for a review of other measures, cf. Borleffs et al. 2017). Their method is based on the premise that for the graphemes (or, more generally, graphematic modules) of different writing systems, there exist different types of correspondence rules. In segmental phonographic writing systems, for instance, a correspondence rule is defined as the relation of a basic shape to the phoneme(s) (in the authors' terminology "phones") that it relates to. Based on their complexity, correspondence rules are assigned different numerical values, the minimal value being 1. If a basic shape relates only to a single phoneme (such as |f| in German), the associated correspondence rule is assigned this minimal numerical value of 1. To measure the transparency of a writing system, the values of all correspondence rules in a grapheme inventory are added up and divided by the number of basic shapes. This results in the value of graphematic transparency, the so-called *gt-value*. The gt-value that Neef & Balestra (2011) calculate for German is based on Neef's (2005) detailed analysis of the graphematics of German, which introduced all relevant correspondence rules. To illustrate how gt-values of different systems may be compared, Neef & Balestra, based on several preliminarily assumed correspondence rules, also calculate the gt-value for the graphematic module of Italian. The German gt-value is 2.05, the one for Italian 1.36. Evidently, the value for Italian is much closer to 1, which is the most natural value of transparency. In sum, thus, the graphematic module of Italian is, on the linguistic parameter of transparency, descriptively more natural than the graphematic module of German.

Neef & Balestra's proposal is a remarkable starting point for a universal method of measuring graphematic transparency. Indeed, the application to other alphabets or writing systems of other phonographic types appears straightforward: for every writing system, be it an alphabet, an abjad, an abugida, or a syllabary, system-specific correspondence rules must first be formulated. A challenge, in this context, could be assigning numerical values to different types of correspondence rules across different types of writing systems. In this vein, a universal repertoire of correspondence rules and numerical values that every (type of) writing system can draw from is a desideratum, and Neef & Balestra (2011: 114) themselves remark that "[w]riting systems can be compared with respect to the subset of types of correspondence rules that are relevant for them and with respect to the quantitative role the different rules play". A further restriction of the recoding model on which this type of gt-calculation is based (cf. Neef 2005) is that it considers only individual segmental units of writing, whereas the transparency of sequences (except for fixed letter combinations, cf. Neef & Balestra 2011: 121f.; Section 2.2.1) such as graphematic syllables, graphematic words, etc. is not incorporated into the overall gt-value of the system. In the present conception of a functional grapholinguistic theory, this question of compositionality is investigated by a different parameter (*compositional transparency*, cf. Section 6.6). Also, it must be noted that due to its restriction to phonographic correspondences, gt-values completely disregard other types of transparency in a writing system; this includes morphographic transparency, which is important also in phonographic writing systems (such as German, cf. Schmidt 2018; Berg 2019). That these different types of transparency are in conflict is discussed in the context of the processing fit (cf. Section 7.2.4).

In the following, it will be sketched how the method of calculating a gt-value can be applied to writing systems that are typologically non-alphabetic and other approaches to evaluating graphematic transparency will be presented.

As already mentioned, abjads such as Arabic and Hebrew should pose no problems for the above-described calculation of a gt-value. The fact that short vowel phonemes are not represented in writing is not a problem since it does not interfere with graphematic transparency at the level of individual graphemes (it does, however, at the level of compositional transparency). As long as individual graphemes are paradigmatically transparent, the gt-value will be natural and indicate overall transparency. A further aspect specific to Arabic is positional allography. Positional allographs of Arabic graphemes are by definition not individual graphemes but variants of one grapheme (cf. Section 2.3). Thus, a single correspondence rule should still suffice for every grapheme; it is not necessary to assume different rules for the different basic shapes that are realized depending on the grapheme's position within a word. The same applies to the allographs <σ/ς> in Greek. Both of these instances of allography, from a graphematic point of view, do not behave differently than, for instance, German <f> that was mentioned above. Of course, unlike Arabic graphemes or Greek <σ/ς>, German <f> does not change its basic shape depending on its position within a larger context (except maybe for sentence-initial capitalization, cf. Section 2.3) but, like Arabic graphemes or <σ/ς>, regardless of its position, it always relates to the same phoneme. Positional allographs of this kind, thus, do not affect graphematic transparency. It is expected, however, that they are relevant for the opposite direction of analysis and the resulting graphematic *uniformity* (or, in Neef's terms, *orthographic transparency*, see below).

An abugida that deserves special attention is Thai, as it is sometimes claimed to be "the most complicated writing system in the world".[224] It poses two unique challenges for the calculation of the gt-value described above. Firstly, it marks tone in a complex and suprasegmental manner. Secondly, while only a restricted number of consonant phonemes (/p, t, k, m, n, ŋ, w, j/) are phonotactically licensed in syllable-final position, graphematically, all consonant graphemes can be written in syllable-final position. For this very reason, basic shapes occurring in syllable-final position 'switch' their graphematic relation: in other words, they cease to relate to the consonant phonemes they prototypically relate to in all other positions and are instead 'reassigned' to correspond with one of the few phonemes phonotactically licensed in that position. This tampers quite severely with transparency, if only in specific positions. These two features of Thai call for a clarifi-

**224** Cf. https://www.youtube.com/watch?v=gKVtpCByEy4 (July 21st, 2020) for a video titled "World's Most Complicated Writing System".

cation of how Neef & Balestra's (2011) method of calculating a gt-value could be modified or extended by additional measures of transparency.

(1) When segments are treated individually, the fact that Thai boasts six basic shapes or graphemes[225] (|ฐฑฒถทธ|) for syllable-initial /$t^h$/ should not pose a significant problem for transparency (but ample problems for uniformity, see below). Each of these shapes is in a graphematic relation with /$t^h$/. However, the resulting graphemes are members of distinct consonant classes (high, middle, and low). Their class membership is one of the multiple factors that contribute to the assignment of tone to written syllables. Thus, in total, the informational load of these consonant graphemes is richer as they are not merely in graphematic relationships with consonant phonemes. Since tone assignment is a suprasegmental affair, however, it cannot be grasped by means of the paradigmatic segmental parameter of transparency. Like some of the other phenomena described above, it must be treated by the parameter of *compositional transparency* instead (see below).

(2) While suprasegmental tone marking does not affect the gt-value as it was presented here, the second problem does. Take the basic shape |ร|: it is, in syllable-initial and syllable-medial position, graphematically related to the phoneme /r/. However, since /r/ is not licensed as a syllable-final phoneme in phonological syllables but |ร| is still written in this position, the graphematic relation of |ร| in syllable-final position is 'redirected' from /r/ to /n/. Thus, as a grapheme, <ร> is not completely transparent and must have a gt-value of (at least) 2 even if the change in graphematic relation is perfectly predictable from the grapheme's position. The situation is analogous for all graphemes that are not in graphematic relations with the phonemes licensed in syllable-final position, i.e. /p, t, k, m, n, ŋ, w, j/. In any case, this feature of the writing system of Thai increases the overall complexity and gt-value of its graphematic module.

Another feature that cannot be captured by the paradigmatic parameter of graphematic transparency is linearity and order: in Thai, the sequence of graphemes does not always straightforwardly correspond to the sequence of phonemes that they individually relate to. This is due to misaligned vowels, i.e. vowel graphemes that are positioned to the left of consonant graphemes although, in the phonological representation, the vowel follows the consonant. In other words, if the vowel is misaligned, a graphematic sequence such as <VC> is in a graphematic relationship with a phonological sequence /CV/. Such idiosyncrasies of writing systems that add positional opacity to the graphematic module are treated by the syntagmatic parameter of *positional transparency* (see below).

Shifting the focus away from different types of phonographic correspondence rules, a relevant question is how correspondence rules can be modeled in morphographic systems such as Chinese. Unsurprisingly, various features of

225 A highly complex question concerning Thai that cannot be answered within the scope of this book is whether the basic shapes that relate to the same phoneme but are assigned to different graphematic consonant classes are allographs of one grapheme or individual graphemes.

Chinese cannot be accounted for by the correspondence rules described by Neef & Balestra (2011).

The first obvious difference is that instead of phonemes, the morphographic graphemes of Chinese relate to morphemes. Notably, if a well-known measure closely associated with graphematic transparency, *orthographic depth* (cf. Section 7.2.4), which is usually interpreted phonologically and employed to distinguished shallow (= transparent) from opaque (= intransparent) phonographic writing systems, is extended to account for morphographic orthographic depth, “Chinese [...] is quite shallow – especially in comparison to alphabetic writing systems” (Handel 2013: 33). Chinese graphemes, evidently, are overwhelmingly transparent.

Since the naturalness parameter of transparency is concerned with graphemes, a question that was already discussed in Section 2.2 and concerned Chinese graphemes must be critically reevaluated: are the “characters” of Chinese (i.e. basic shapes occupying segmental spaces) graphemes even if they are complexly structured? Or are subsegmental components – both phonological and semantic ones – graphemes? The earlier stance was that subsegmental components are undeniably functional in that they correspond with linguistic information of some sort, whether a semantic field or a (in most cases approximate) pronunciation. However, their correspondences are not stable and they do not relate to linguistic *units*. Thus, they are parts of complex graphemes, which, like simple graphemes, relate to morphemes. At the holistic level of graphemes, now, correspondence rules are, as mentioned above, almost maximally transparent: each basic shape (with very few exceptions) relates to one morpheme. For example, a basic shape such as |妈| relates unambiguously to the morpheme {mother} with the phonological representation *mā* in Mandarin (and different pronunciations in other varieties of Chinese). In such a holistic analysis, the graphematic relations of possible subsegmental components are disregarded, in this example the radical for “female” and the phonological component |马| deriving originally from *mǎ* ‘horse’.

However, it is possible to evaluate graphematic transparency also at this subsegmental level when the correspondences of every semantic and phonological component are assessed. If this is done, a relevant question can be answered: are the subsegmental components as transparent as the complex graphemes they constitute? They are not. Both types of components – semantic as well as phonological – are not completely transparent (which is also one of the reasons they are not treated as graphemes). Consequently, an evaluation of graphematic transparency that accounts for the transparency of subsegmental components as well as the compositional values of the complex graphemes they constitute will result in a much higher (and thus less natural) gt-value for the graphematic module of Chinese than a holistic analysis that does not break up complex graphemes into their components.

To make possible a comparison of different grapheme-based gt-values, from a descriptive point of view, the holistic method is preferable. This, however, does not necessarily reflect psychological reality, since, in processing, subsegmental components are relevant entities (cf. Section 7.2.6). Notably, evaluating graphematic transparency at a subsegmental level is possible also for other writing systems: in German, for example, or any other writing system employing Roman script, one could analyze how transparently the subcomponents of basic shapes represent phonological features (cf. Primus' analyses, Primus 2004, 2006). The same is true for the correspondence (or non-correspondence) of visual features and linguistic features in Thai (cf. Section 5.1).

An additional problem faced by an evaluation of transparency that incorporates phonological components is their applicability across varieties of Chinese. In fact, they give (more or less) reliable pronunciation clues only in Mandarin while other varieties such as Cantonese cannot rely on them. This is relativized by the fact that children of all varieties are instructed in Mandarin in school (cf. Anderson et al. 2003). Although the "same writing system" is allegedly used across varieties, at the subsegmental level, it is not actually the same system. This is not surprising since, in the modular model of writing systems (cf. the introduction to Part II), language systems are defined as base modules of writing systems. In the case of Chinese, a number of mutually unintelligible (spoken) varieties simultaneously serve as base modules of a unified morphographic system. The graphematic modules of the Mandarin, Cantonese, etc. varieties are fundamentally different with respect to the transparency of subsegmental phonological components and, in general, the pronunciation of morphemes that graphemes relate to. The only stable constant is the semantic components that keep their values across varieties. "The Chinese writing system", if seen in this constrained monolingual view, is, thus, at a subsegmental level only suited for Mandarin and not for other Chinese varieties. This makes it likely that the sociocultural fit of the Chinese writing system is drastically reduced for users of varieties other than Mandarin (cf. Section 8.3).

An interesting question that concerns the subsegmental level in Chinese is its interaction with the segmental level: while the evaluation of the transparencies of both levels has been mentioned, what about the question whether the components, when combined to form graphemes, contribute to its overall graphematic value in a transparent manner? This question is treated by the separate parameter of *compositional transparency* (see below). It is the analog of the parameter of morphosemantic transparency in Natural Morphology (cf. Section 4.2.3) that identified compounds such as *blueberry* as more natural than *cranberry*, since in *blueberry*, the meanings of the components contribute transparently to the overall meaning of the word, which is not the case in *cranberry*. It is also noteworthy that this type of transparency represents a form of diagrammaticity: if a specific semantic component always transparently indicates the same meaning, and consequently, the complex graphemes that feature it are members of one semantic field, then this semantic component establishes paradigmatic endophoric diagrammaticity (cf. Section 6.2).

The inclusion of the parameter of compositional transparency is paramount as it can be used to evaluate not only the structure of complex Chinese graphemes in Chinese but generally of larger graphematic units, including graphematic syllables, graphematic words, etc. The parameter of transparency alone as described in this section evaluates exclusively graphemes, and it does so individually and paradigmatically. Thus, as was highlighted in the context of complex tone marking in Thai, no suprasegmental, contextual graphematic operations can be grasped by this type of transparency. Taken alone, it is not informative: while the graphemes of Thai may individually be transparent, compositionally and positionally, the graphematic module of Thai boasts a high degree of unnatural features. To disregard these latter dimensions means disregarding crucial factors that contribute to the overall linguistic fit of a graphematic module.

To sum up, following Neef & Balestra (2011), the transparency of a writing system's graphematic module can be evaluated by formulating correspondence rules for all its graphemes. Graphematic relations that are not transparent, i.e. relations in which a basic shape relates to more than one linguistic unit, are assigned higher numerical values than unambiguous relations. In the end, the values of correspondence rules are added up and divided through the number of basic shapes. The result – from the smallest and most natural gt-value of 1 upwards – indicates the segmental graphematic transparency of a writing system. It is segmental insofar as compositional and positional transparency, which are affected by suprasegmental graphematic features, are not accounted for. As shown above, it is possible to transfer this method of calculation to other types of writing systems as well, including morphographic writing systems.

## 6.5 Uniformity

The parameter of uniformity is complementary to transparency and thus evaluates how many visual units linguistic units correspond with. As established in Section 2.2, this perspective is characteristic of the referentialist view of graphematics in which a language's phoneme inventory serves as a starting point and the main goal is to evaluate how the individual phonemes can be written (cf., for example, Garbe 1985 for a list of how German phonemes are represented in writing). While Neef & Balestra (2011) refer to transparency (as described above) as *graphematic transparency*, they term uniformity *orthographic transparency*. This is based on their argument that graphematic transparency affects decoding, i.e. reading, whereas orthographic transparency has a bearing on production processes, i.e. writing, and particularly writing in an orthographically correct manner. Accordingly, the designation *orthographic* transparency is not accidental: in German, for which the phenomenon was described, spellings are not merely graphematic, but always judged against an orthographic norm. The orthographic module is superimposed upon the graphematic module, and even when the orthographic module licenses as correct a spelling that is located inside the graphematic solution space, this spelling is still phenomenologically an orthographic spelling rather than a graphematic variant. The question of whether it is true that uniformity influences spelling will be discussed in the context of its processing fit (cf. Section 7.2.5).

A descriptive question must be answered before uniformity can be reasonably evaluated for a given writing system: which type of linguistic unit its written units correspond with. This question is at the heart of writing system typology (cf. Section 2.7). It is not an entirely trivial question since no writing system can be classified as being purely of a single type of writing system. The mixing of different types of graphematic relations within a single writing system leads to conflicts of uniformity between different linguistic levels, e.g. phonology and morphology. In German, for example, a degree of morphography adds to the uniformity of how morphemes are graphematically represented, which simultaneously diminishes the uniformity of phoneme representation. For example, the orthographic principle commonly known as *morpheme constancy* overrides a transparent representation of devoiced consonants in final position: accordingly, while in the plural form <Hunde> 'dogs', the <d> actually stands for /d/, in the singular form <Hund> 'dog', word-final <d> is written although here, on account of final devoicing, it relates to /t/. In this case, a lack of transparency results in a lack of phonographic uniformity: /t/ can be written as <t> in all contexts within a syllable, but syllable-finally, it additionally corresponds with <d>. The fact that the uniformity of the graphemic representation of /t/ is disrupted only in this specific context (syllable-finally) ameliorates the situation somewhat. This becomes evident when other cases such as /f/ are considered, which is represented ambiguously as <f>, <v>, or <ph> independently of the context within a word (but dependent on several other factors).

Consequently, writing systems that mix phonography and morphography have two separate values of uniformity: a value of *phonographic uniformity* that assesses how uniformly phonemes are 'represented' by graphemes as well as a value of *morphographic uniformity* that evaluates how uniformly morphemes and morphonological alternations are 'represented'. In a language with only few phonological processes that alter the phonological representation of morphemes (such as Finnish) phonographic and morphographic uniformity converge and the two principles are not in conflict. When the phonological representation of a morpheme is not consistent, however, a writing system's graphematic module must choose whether to stay faithful to phonology or morphology. Notably, this conflict between the two types of uniformity is relevant only in writing systems that can, in some way, incorporate both the phonological and morphological levels. This is most clearly the case for segmental phonographic writing systems. It is trivial to observe that when the unit of representation in a writing system is the phoneme, then morphemes made up of phonemes can also be represented in writing and the relationship between 'dominant' phonography and 'optional' morphography can always be assessed.

When the unit of representation is the morpheme, however, as in morphographic Chinese, phonology is either not represented at all or represented only partially. This goes to show that, following the *strict layer hypothesis*, the representation of lower graphematic levels always implies the representation of higher graphematic levels. When the lowest represented level in a system is the morpheme level, however, the representation of levels beneath it is optional and commonly not consistent (cf. Table 5 in Section 2.7). In the writing system of Chinese, there are approximately 800 phonological components for roughly 400 phonological syllables of Modern Standard Mandarin (cf. Handel 2013: 50). This means that syllables are not uniformly represented by the subsegmental phonological components in graphemes. An analogous analysis of subsegmental uniformity can also be carried out for the semantic components; here, the main question would read: do morphemes that are part of one semantic field uniformly contain the same semantic component? Notably, while in Chinese, phonography and morphography are in conflict (albeit only to a small degree), this does not apply to the morphographic part of the Japanese writing system. In Japanese, morphographic graphemes (*kanji*) do not contain phonographic elements, which means that the only reasonable question that can be investigated here is whether the morphemes of Japanese are uniformly represented by kanji graphemes.[226]

A feature of writing systems that conflicts with uniformity (as already foreshadowed in the context of transparency) is positional allography. For example, in Arabic, most consonant phonemes are represented by graphemes that are materialized by four distinct basic shapes. These alternate depending on whether a grapheme occurs in isolation or in sequence; if the grapheme is part of a sequence, what is important is its position (initial, medial, final). Consequently, consonant phonemes are not uniformly represented by single basic shapes but by a set of basic shapes (exhibiting, to varying degrees, visual similarity). Notably, since positional allographs are per definition different variants of a single grapheme, phonemes are technically still uniformly represented by abstract graphemes. The question is whether the existence of alternative basic shapes that lack a distinct graphematic status results in unnaturalness on the parameter of graphematic uniformity. Above, it was established that for transparency, positional allography is not unnatural because the existence of multiple basic shapes that are assigned to the same grapheme and thus relate to a single linguistic unit only increases the number of units; these units, however, are themselves transparent. In the case of uniformity, the opposite is the case: a given linguistic unit corresponds with multiple basic shapes. Again, descriptively, it can be argued that these basic shapes are assigned to the same grapheme, which, however, ignores their distinctiveness[227] at the graphetic level. For this reason, I argue that positional allography reduces

**226** A degree of complexity is added, however, because Japanese morphemes often have several possible readings, i.e. pronunciations. This, however, is not primarily a graphematic but a morphological problem.

**227** Consider, however, the counterargument that positional allographs, at least in Arabic, visually resemble each other more than basic shapes that are part of different graphematic relations (i.e. relate to different consonant phonemes). But consider also <σ> and <ς> in Greek, for which this is evidently not the case.

uniformity. In any case, what effect this has on the systems that exhibit positional allography must be assessed in system-specific analyses.

To give another example, take Thai. Here, ambiguity is rampant, and accordingly, uniformity is low. As discussed in the preceding and following sections, the writing system of Thai exhibits graphematic unnaturalness with respect to a range of semiotic naturalness parameters. It most clearly does so with respect to uniformity, however. In a mere paradigmatic analysis that ignores positional constraints and effects of context-dependency, there are, for example, a whopping seventeen possibilities of writing /t/. This is partially caused by the phonological peculiarity of Thai that phonotactically, in syllable-final position, the only allowed consonant phonemes are /p, t, k, m, n, ŋ, w, j/ (see above).

While there exist systems that are almost biunique (such as Finnish), i.e. exhibit the most natural values on the parameters of transparency as well as uniformity, most writing systems exhibit naturalness on only one of these two parameters. Modern Greek and French, for example, have fairly transparent graphematic modules, which, however, are not uniform given that many of the languages' phonemes do not uniformly correspond with a single grapheme.

## 6.6 Compositional transparency

Compositional transparency is concerned with the question of whether units (or segments) at a given level add up to constitute the value or meaning of units at a higher graphematic level; for example, whether the graphematic values or functions of a grapheme's elementary forms transparently add up to the overall value or function of a grapheme. Compositional transparency can be evaluated at the subsegmental, segmental, and higher graphematic levels. Notably, there are various phenomena that result in compositional opacity: (1) *graphematic excess* is when the graphematic string provides more information than would be necessary for a graphematic representation of the corresponding linguistic unit; (2) *graphematic underspecification* occurs when not all parts of the signatum are represented in the signans, and (3) *graphematic mismatch* is when the values of an element's parts do not add up to the value of the element as a whole.

A central context for the evaluation of compositional transparency is complex graphemes. Complex graphemes are graphemes that are graphematically but not graphetically segmental (cf. Section 2.2). An example is German <ch>: it consists of two basic shapes but is a single complex grapheme. From the perspective of compositional transparency, <ch> represents a graphematic mismatch since its relation to the phoneme /x/ does not follow compositionally from the graphematic functions of its two constituents |c|, which itself is not a grapheme, and <h>.[228] Another important context for compositional transparency is graph-

228 It is justified to ask what exactly |c|, which itself is not a grapheme, should contribute to the overall "meaning" (= signatum) of the complex grapheme if its non-graphemat-

eme combinations: in Thai, the vowel grapheme <◌ะ>,[229] which is bound and thus always attached to a consonant grapheme, is in a graphematic relationship with the short vowel /a/. However, it is also part of several grapheme combinations in which it 'loses' this value: the complex combination <เ◌ะ> relates to the short vowel /e/, for example, and the combination <แ◌ะ> relates to the short vowel /ɛ/. These are not complex graphemes since both parts are independent graphemes, respectively: <เ◌> relates to /eː/, <แ◌> corresponds with /ɛː/. In the combinations <แ◌ะ> and <เ◌ะ>, thus, <ะ> marks the shortness of the vowel and does not correspond with /a/, which it usually does. Compositionally, these complex vowel grapheme combinations relating to single vowel phonemes (instead of diphthongs, for example) result in compositional opacity.

In general, what interferes with compositional transparency is when certain linguistic units remain graphematically unrepresented. In the segmental phonographic writing system of Arabic, short vowel phonemes are commonly not represented graphematically, leading to graphematic underspecification. A sequence such as <كتب> is, context-independently, decoded as /ktb/. While as a triconsonantal root, this bit of phonological representation is undeniably the most relevant information to be represented graphematically, from the perspective of compositional transparency, it is ambiguous. It serves as the basis for many words, among them a slate of words in which all vowels are short and which are, thus, all written as <كتب> /ktb/: examples are /kutub/ 'books (plural)', /kataba/ 'he wrote', and /kutiba/ 'it was written (masculine)'. For the paradigmatic parameters of transparency and uniformity, this absence of short vowels in the graphematic sequence is not of relevance. However, syntagmatically, compositionality is disturbed because a graphematic representation devoid of short vowel graphemes does not add up to the phonological representations of corresponding words.

Another example of compositional transparency concerns complex tone marking in Thai (cf. Figure 32). It is determined by (1) the type of the syllable, i.e. whether a syllable is *alive* (ending in a long vowel or in /m, n, ŋ, w, j/) or *dead* (ending in a short vowel or in /p, t, k/), (2) the class of the syllable-initial consonant grapheme (*high, middle, low class*), and (3) the quantity of the vowel (*long, short*) (cf. Smyth 2002: 16). These three variables interact in complex ways to mark five tones: low, mid, high, rising, and falling tone. The existence of four tone markers in the writing system of Thai does only little to reduce the complexity of

ic status means it does not actually relate to a linguistic unit. Consider the trema in German <ä>, <ö>, and <ü>. It is a diacritic and it modifies the graphemes <a>, <o>, and <u> in a systematic way, signaling fronting of the vowels. It is not itself a grapheme since it is not in a graphematic relation with any linguistic unit. Unlike |c|, however, it does not occupy its own segmental space and it is not part of the same class of basic shapes as |a, o, u|. It would still be imaginable that a graphetically segmental unit such as |c| functions like the trema, i.e. modifies the grapheme it combines with in a systematic way.

229 In the notation of Thai vowel graphemes, the dotted circle |◌| is a placeholder for a given consonant grapheme that the vowel grapheme depends on/attaches to.

the situation, as they are transparent only to a limited degree and still interact with the mentioned three variables in complex ways. For example, in combination with consonant graphemes of the high and middle classes, the low tone marker actually marks low tone, whereas, in combination with consonant graphemes from the low class, it marks falling tone. When combined, these three – with the inclusion of tone markers four – variables represent tone transparently, i.e. are an example of compositional transparency. By contrast, individually, these graphematic resources of written Thai are not transparent. Inversely, the situation is even more unnatural, as the graphematic representation of the five tones is not even remotely uniform.

| | no tone mark | | | | tone mark | tone mark | tone mark | tone mark |
|---|---|---|---|---|---|---|---|---|
| | $-V_L(C_N)$ | $-V_S(C_N)$ | $-V_L C_S$ | $-V_S C_S$ | ่ | ้ | ๊ | ๋ |
| low class | คีน medium | คิน high | คีจ falling | คิจ high | ค่ falling | ค้ high | | |
| medium class | กีน medium | กิน medium | กีจ low | กิจ low | ก่ low | ก้ falling | ก๊ high | ก๋ rising |
| high class | ขีน rising | ขิน low | ขีจ low | ขิจ low | ข่ low | ข้ falling | | |

Figure 32. Complex graphematic representation of tone in Thai, adapted from https://gte-localize.com/wp-content/uploads/2018/06/thai1.gif (February 9th, 2019; page is not accessible anymore)

A straightforward application of compositional transparency comes in the form of complex Chinese graphemes. A complex grapheme consisting of a semantic and a phonological component is considered compositionally transparent if both components are maximally transparent themselves, i.e. if the phonological component reliably represents the pronunciation of the morpheme (that the entire grapheme relates to) and the semantic component correctly indicates the semantic class membership of said morpheme. Similarly, in complex graphemes that consist of two or more semantic components, natural values of compositional transparency are achieved when all components contribute semantically to (a graphematic representation of) the morpheme's overall meaning.

## 6.7 Positional transparency

The syntagmatic parameter of positional transparency, similar to morphotactic transparency, the natural morphological parameter it is modeled after (cf. Section 4.2.3), is concerned with whether the sequences of elements in the signans and the signatum are congruous. If, for example, in the signatum, e.g. the phonological representation of a graphematic word, the sequence is /abc/, then what we would expect to be the default is that the corresponding graphemes in the signans appear in that same sequence: <abc>. Notably, this is not always the case.

One of the most prominent examples of unnaturalness with respect to positional transparency is misaligned vowels in Thai (see above). In essence, some vowel graphemes completely or partially *precede* consonant graphemes despite the fact that in the phonological representation of the graphematic words in question, the corresponding vowel phonemes *follow* the consonant phonemes: take the graphematic word <แปลก> <ɛ:plk> *plek* 'strange'. It is decoded as /plɛ:k/ although the vowel grapheme <แ◌> precedes the graphematic consonant cluster <ปล> <pl> in the written form of the word (cf. Winskel 2009: 22). From a linguistic (or more generally, semiotic) point of view, this is unnatural, and an obvious assumption is that it also leads to problems in processing (cf. Section 7.2.7).

Commonly, in larger graphematic contexts such as the graphematic word, it is a trivial matter to determine the position of a grapheme: in <has>, <h> is self-evidently word-initial, <a> is word-medial, and <s> is word-final. Logically, if isolated from a context and presented individually, it is not (straightforwardly)[230] possible to determine the position in which a grapheme was produced. It is possible in Arabic, however, as well as for Greek <σ> and <ς>. In other words, positional allography (cf. Section 2.3) adds visual information about a grapheme's position: due to the appearance of a specific allograph, it is clear where in a graphematic word a grapheme is/was positioned. (This, to circle back to what was discussed above, does not reveal anything about whether the positions of phonemes in the phonological representation and corresponding graphemes in the graphematic word are congruous, however.) Descriptively, the additional positional information provided by positional allography, as is the case in Arabic, is not needed, pointing to the fact that it has developed productionally (cf. Section 7.1.2). It leads to naturalness conflicts with other parameters (such as uniformity, see above) and generally violates the principle of economy as it provides information that is more or less (again, from a descriptive perspective) redundant.

## 6.8 Figure—ground

Figure—ground is different than the other parameters. As established in the framework of Natural Morphology, it is a perceptual parameter (cf. Section 4.2.3). It can be analyzed from two different perspectives: on the one hand, (1) the perspective of (parts of) the signatum and signans within a sign as well as the relations of these parts with each other, which corresponds with the perspective taken by the other parameters described above; on the other hand, (2) the perspective of the entire sign and its relation with other signs. In total, a fine-grained analysis distin-

230 Notably, in cursive handwriting, due to effects of coarticulation, even an isolated grapheme (or rather the concrete graph in which it is materialized) might provide clues about the position in which it was initially produced.

guishes three (and possibly more) levels: the *subgraphemic* level, at which parts of graphemes and their relations with each other are investigated; the *graphemic* level, studying graphemes and their relations with each other, and the *supragraphemic* level, where larger graphematic units (such as graphematic words or sentences) and their relations with each other are analyzed. All three will be treated exemplarily below. What must be said in advance is that regardless of the level at which it is analyzed, the basis of figure—ground is perceptual, specifically perceptual salience: one element (whether a sign or part of a sign) is perceptually more salient than another element (a sign or part of a sign), with the figure being the more salient and the ground the less salient element.

SUBGRAPHEMIC LEVEL. The first level at which figure—ground relations can be evaluated is *within* signs. If visual salience is to be evaluated at this level, the visual constituent of the sign in question, its basic shape, must be segmented into smaller parts. An example of a rather straightforward segmentation that has already been mentioned concerns Chinese graphemes. Here, the different components in a basic shape/grapheme can be analyzed with respect to figure—ground; necessary information that makes this possible is provided top-down by higher levels (i.e. the graphematic level and specifically an analysis of the functions the components fulfill). If graphemes consist of two components, which is the case for many of them, these components are equivalent with respect to figure—ground if they take up an equal amount of space within the segmental space. By contrast, if graphemes consist of three, four, or five components, these are often not of the same size and thus do not occupy the same amount of the segmental space available for the entire grapheme.[231] If size is to be taken as a decisive criterion for the evaluation of visual salience, which is contestable,[232] then smaller components serve as grounds and larger components as figures. Figure 33 offers a few examples of graphemes consisting of three or four components; examples of five-component graphemes are given in Figure 34.

231 Note that there are exceptions: in several three-component and four-component graphemes, components take up an equal amount of the segmental space. Examples include <術> *shù* 'technique, art, skill', where the components each take up a third of the space, or <鬆> *sōng* 'loosen, release', in which the four components each occupy a subsquare of the segmental space, which is divided horizontally and vertically in the middle. Note that these two graphemes are traditional and not simplified graphemes (cf. Takagi 2014: 88f.; also Section 5.2).

232 Other features that could also be considered as measures for visual salience and operationalized for an evaluation of figure—ground relations are the visual complexity of the components or the position of the components within the segmental space, with some positions possibly being more visually salient than others.

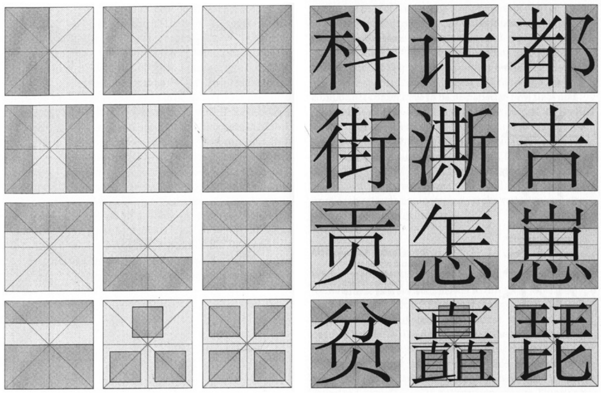



FIGURE 33: Subdivision of the segmental space in Chinese script, from Palmer (2015: 32–33)

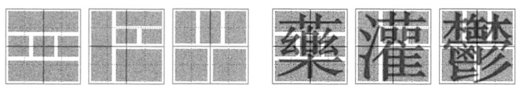



FIGURE 34. Five-component basic shapes, adapted from Takagi (2014: 88–89)

Up until this point, figure—ground was only analyzed graphetically since only parts of the signantia of Chinese graphemes were compared with each other. In the next step, it could be analyzed whether graphetic figure—ground relations correspond with figure—ground relations in the signatum. For example, whether, in a grapheme consisting of three semantic components, the component that is visually most salient (because it is the largest) somehow contributes more or even most of the meaning to the morpheme that the whole grapheme relates to. If this were the case, we would have also identified a graphematic figure—ground distinction in the grapheme.

GRAPHEMIC LEVEL. The first and most relevant manifestation of a figure—ground distinction at the graphemic level is spacing. Empty spaces, despite not being graphemes themselves, make visible written units of various sizes by setting them apart from or contrasting them with an empty background. On this very page, empty spaces are white (or whitish); these spaces of 'nothingness' bring order to the text both graphetically and graphematically. Figure—ground relations are more natural when visual units that are functionalized graphematically (to relate with different types of linguistic units) are set apart both from the background and each other by spacing. Note that different graphematic units can also be set apart by other measures such as the alternation of scripts in the writing system of Japanese, which indicates word boundaries in the absence of spaces between words. This is arguably also natural in that it allows readers to identify word boundaries; however, script alternation and the lack of spaces is arguably not equivalent to the presence of empty spaces when it comes to the parameter of figure—ground as empty spaces are visually more salient. At this point, it is important to note that from the perspective of processing, spacing is one of the most important and well-studied graphematic phenomena (cf. Section 7.2.8).

One of the first examples that come to mind regarding the varying salience of graphemes is the difference between vowel vs. consonant graphemes in (many) abjads and abugidas. For example, in abjadic Arabic, the consonant graphemes relating to short vowels, which are optional, are visually less salient than their consonant counterparts: they are smaller and also do not occupy their own segmental space but are instead positioned (like 'graphematic clitics') above or below consonant graphemes. In Thai, typologically an abugida, vowel graphemes are not optional. However, with some exceptions, their visual features are the same as for Arabic short vowel graphemes: they are mostly smaller in size and often positioned below or above consonant graphemes. A remarkable difference between Arabic and Thai is that some Thai vowel graphemes are as salient as consonant graphemes; they are equal in size, occupy their own segmental space and 'stand' on the base line of the linear space. That there are secondary vowel graphemes that are smaller but also vowel graphemes that behave like consonant graphemes results in a decrease of graphematic naturalness on a range of parameters: diagrammaticity, compositional transparency, and positional transparency, and, of course, figure—ground.

Thus, in Arabic and, for the most part, also Thai, vowel graphemes are the grounds to consonant graphemes, which are the figures. For Arabic, this makes perfect sense: it has been elaborated abundantly in the literature that Arabic makes use of consonantal roots (cf. Ryding 2014: 61–63). By contrast, for Thai, it would also be imaginable that consonant and vowel graphemes were equivalent since nothing in the structure of Thai phonology or morphology specifically renders consonants more salient than vowels. This is also echoed by the fact that the writing system of Thai is treated as an *alphabet* rather than as an abugida in several works (cf., for example, Winskel 2009), a classification that neglects that Thai differs from true alphabets in important structural respects. Other conceptions acknowledge that Thai is different from alphabets structurally but highlight the fact that for processing, users treat it like an alphabet (cf. Rimzhim, Katz & Fowler 2014), suggesting a discrepancy between structural description and cognitive reality/use.

A difference between the salience of consonant and vowel graphemes can also be attested in Roman script and alphabets that employ it as well as the Greek, Georgian, and Armenian scripts and writing systems: as discussed in Section 2.4, basic shapes that are being used for consonant graphemes in these writing systems frequently exhibit the feature [+length] while the compact basic shapes that relate to vowel graphemes do not. Consonant graphemes, thus, are for the most part visually more salient in these systems.

The above-mentioned optional short vowel graphemes of Arabic are independent graphemes and – when they are included, for example in children's books or L2 teaching material – serve as parts of the graphematic word. From a visual point of view, there exist similar units in other writing systems: *diacritics*. They are, however, only parts of graphemes instead of independent graphemes. Take Roman script: it has been adopted for a multitude of alphabets and has thus witnessed a large degree of modification. A common strategy in this context is adding diacritics to existing basic shapes to form new graphemes (cf. Daniels 2006a: 17–20). These diacritics – for instance, the cedilla in <ç>, the háček in <č>, or the tilde in <ñ> – take up only part of the segmental space and are enclitic to the basic shapes/graphemes they become a part of. They are arguably the ground to the basic shapes as their figures. Commonly, diacritics modify not only the basic shape they are added to. They simultaneously also alter the associated graphematic relation in a systematic, diagrammatic way. In the Czech alphabet, for example, the háček indicates postalveolar articulation as in <č>, <š>, or <ž>, or palatalization as in <ň> or <ě>. Note that as bound units dependent on other units, diacritics, unlike vowel graphemes in Arabic and Thai, must be analyzed from a subgraphemic perspective. Graphetically, these two phenomena are equivalent but graphematically, they are different.

So far, only 'scriptual' basic shapes, i.e. units of a given script that is employed by a given writing system were mentioned. However, other types of basic shapes (cf. Section 1.2.1) are also an important part of writing systems. Take punctuation: it has often been noted that punctuation marks, especially the most frequently used ones, the period and the comma, are visually less salient than the graphetic material – mostly scriptual basic shapes – surrounding them. Reasons are their small size and the associated small amount of segmental space that they occupy (as well as their position inside the segmental space). Undeniably, thus, they can be conceived of as grounds to the other types of basic shapes. Notably, their lack of visual salience has a significant effect on how these punctuation marks are processed (cf. Hill & Murray 2000 for an analysis of how the comma influences processing).

Supragraphemic level. At the highest graphematic level, figure—ground, as described by Dressler (1989) in the framework of Natural Textlinguistics, can distinguish between elements of a text such as footnotes, headlines, and the running text, categories that are characterized mainly by the functions they fulfill in the larger context of the entire text but that are at the same time commonly also visually distinct. Certain types of graphetic variation, including bold type, italics, and color can also serve as instruments that render an element the figure and set it apart from a less salient ground. A prerequisite for this figure—ground distinction to work is the notion of a graphetic default. A word set in **bold**, for instance, can only be perceived as the figure if there is non-bold text in its proximity. For a functional figure—ground distinction, likely an even larger context is necessary: for bold highlighting to be functionalized at the graphematic level (e.g. to be recognized by readers as having the function of highlighting important concepts in the text), at least two more words are needed that are not printed in bold. In this case, for 'statistic' reasons, the bold word – which is in this case also visually more salient – would be perceived as variation, as 'deviance' from the non-bold default (cf. also Meletis 2015: 149f.; Spitzmüller 2013: 126). Notably, visual salience is not always necessarily congruous with a graphematic figure—ground distinction: take a paragraph in which all words except for one word are printed in bold. Graphematically, "non-bold" will be interpreted as deviance from the default, as the figure to the ground (= the rest of the bold paragraph) even though it is, technically, at the graphetic level, visually less salient due to more narrow stroke width, etc. Conceptually, by contrast, the non-bold word is perceptually more salient because it is likely that its 'otherness' immediately catches the eye of readers.[233]

**233** However, it is also likely that a single bold word in an otherwise non-bold paragraph is detected more quickly by readers than a non-bold word in an otherwise bold paragraph. Thus, visual salience and functional salience *do* interact in complex ways. This, of course, is only an assumption and would need to be tested empirically.

# 7 Processing fit

The systematic and linguistic fits are descriptive, providing answers to the questions of how systematic writing systems are and how well they relate semiotically to their languages. The core of explanation, however, is not the systems' structures themselves but how they are used. An important part of this use is processing, i.e. how humans write and read, and more specifically, what goes on with their bodies and in their minds when they write and read. Consequently, physiology and cognition come to the forefront. How they work and, importantly, how they are restrained affects the makeup of writing systems significantly, as writing systems are not only human inventions in the first place but are also communication systems that are continuously used, making them subject to change. In other words, diachronically, the human pressure inherent in processing (and communication) shapes the structures we find in writing systems (cf. Dehaene 2009). This relationship is reciprocal, however: at each given point in time, the structures of writing systems also affect processing.

These considerations constitute the basis of the processing fit that is presented in this chapter. Its subsections deal with each of the three modules of writing systems separately as they all interact with processing in quite distinct ways: Section 7.1 addresses the graphetic processing fit, which focuses on how the material aspects of scripts – i.e. which movements are involved in their production or how they look once they have been produced – have a bearing on our hands, eyes, and brains. Then, Section 7.2 shifts the focus onto the graphematic processing fit. Here, the parameters that were introduced in the context of the linguistic fit will be reevaluated through an investigation of how they interact with human cognition. Finally, a very brief example in Section 7.3 illustrates how orthographic regulation can potentially interact with processing.

## 7.1 Graphetics

> 'Why is this system this way instead of some other way?' Since the alphabet is preeminently a largely undeliberated product of the human mind any profound answer to such a question must obviously take the form, 'What is it about the human mind that has shaped the alphabet to take the form it has?' (Watt 1988: 231)

Scripts are man-made artifacts. They were conceived by humans with the intention of being used by humans. It is, in fact, the continuous use by humans that constitutes the dynamic interaction between human processing needs and features of scripts: their features affect how humans process scripts and, simultaneously, through the conditions of writing and reading processes, scripts are subject to human pressure – the 'human mind', as Watt puts it –, which results in gradual changes of their features. Simply put, the diachronic development of scripts depends crucially on how they are processed by humans. In this section, I will discuss those features of scripts that are particularly relevant in production and perception processes. I will address particularly how features that are not well-suited for processing lead not only to diachronic change but also to challenges in emergent literacy acquisition and argue that this points to fundamental categories of a theory of writing.

Watt (1983b) assumes four main forces to explain why scripts change: (1) *homogenization*, (2) *facilitation*, (3) *heterogenization*, and (4) *inertia*. Crucially, the existence of each feature of scripts to be discussed in this section can be traced back to one of those forces. For Watt, the first two of them are the so-called stronger forces: homogenization results in scripts becoming more homogenous as their basic shapes grow increasingly similar and is motivated by the strive for economy and systematicity. These, in turn, are driven mainly by cognition, more specifically our brain, which "forms systems in the first place" (Watt 1979: 31). In other words, the force of homogenization strives for an optimal systematic fit (cf. Chapter 5) which is achieved by featural uniformity among the basic shapes of a script. As a general cognitive force, homogenization is neither predominantly productional nor perceptual, although Watt (1988: 201) posits, in line with his argument that perception is primary (see below), that homogenization is mainly perceptual since perception is not only a crucial part of cognition but also inherent in production, which always features perceptual feedback loops. Facilitation, as the second major force, is driven by production as it makes shapes easier to produce. Broadly put, facilitation strives for production processes that involve less effort. The two remaining forces, which Watt categorizes as weaker, are perceptual and counterbalance the major forces: heterogenization is the opposite of homogenization and ensures that basic shapes do not become too similar visually to distinguish. Similarity and distinctiveness, thus, are conflicting features of basic shapes as they are constituted by opposing forces. Finally, inertia, as the most passive of the four forces, represents the human preference of retaining the status quo. Notably, it is a fundamental force since it reflects that users generally disprefer and resist changes in the system. This is echoed by a conservatism inherent in the development of writing and the often-described fact that it lags behind the development of speech. Accordingly, the fact that basic shapes "of many scripts have changed strikingly little over thousands of years" (Treiman & Kessler 2014: 159) is a result of the force of inertia. The dynamics and interaction between these four forces of change will be of relevance in the discussion of individual features and their origin.

While graphetic features of scripts can be gathered by a study of existing research on the materiality of writing, an existing list of features can also serve as a promising starting point and as orientation. Such a list is provided in

Treiman & Kessler's (2014) impressively comprehensive account of literacy acquisition. They describe "graphic forms" (Treiman & Kessler 2014: Chapter 5) and what they term the general "surface properties of writing", i.e. the features of the graphetic module of writing systems. These features include artificiality, two-dimensionality, lack of iconicity (which is closely connected to arbitrariness and abstractness) and visual blandness, sequentiality and alternation (as in little internal repetition), rectilinearity, discreteness, and finiteness of writing and its units. In the next step, Treiman & Kessler (2014: Chapter 8) discuss the individual "symbol shapes" and – similar to Watt – the "principles that underlie systems of symbol shapes" (Treiman & Kessler 2014: 153). These are principles that underlie the makeup of scripts, such as economy, conservatism, beauty, expressiveness, similarity, and redundancy. In describing how literacy acquisition is affected by these principles, Treiman & Kessler indirectly emphasize their general relevance for a theory of writing. Thus, their systematic collection of features and principles lends the following section its basic structure.

## 7.1.1 Graphetic features relevant in processing

### Artificiality, arbitrariness, abstractness, and visual blandness

As mentioned above, scripts are artifacts, and accordingly, the first feature of scripts to be discussed here is their *artificiality*. The fact that writing is not intrinsic to a given surface, unlike, for example, the stripes of a zebra, which are an inherent part of the zebra, helps children to grasp the artificiality of writing (cf. Treiman & Kessler 2014: 105). In general, children can observe that writing does not commonly appear on natural surfaces such as cats or leaves. Furthermore, in their environment, children in literate societies have the opportunity to witness the production of writing, e.g. when they observe their parents writing shopping lists. And before the age of 2, children start using verbs such as *make* or *write* when referring to the writing process (cf. Robins et al. 2012), indicating that they have acquired an understanding that writing is man-made. Features of writing closely related to its artificiality are the arbitrariness and abstractness of its shapes. These, in turn, are closely associated with what could be called the "visual blandness" of writing, in other words, the fact that it should not "shout", it should not draw attention to itself. Instead, writing is meant to become invisible behind the linguistic content it conveys to sacrifice the material substance for linguistic meaning (cf. Strätling & Witte 2006: 8). Thus, the graphetic module is relegated to the background, allowing the functions of the graphematic module to come to the forefront. Note that the visual blandness of writing also distinguishes it markedly from drawing.

In this context, arbitrariness equals non-iconicity. In the case of writing, this equals a lack of pictography, which has been termed *non-pictoriality* (cf.

Lavine 1977: 90).[234] Indeed, today, the scripts of the world are largely non-pictorial. Note that iconicity, here, is not meant in a linguistic sense (in which it was discussed in Section 6.2). At the graphetic level of analysis, the fact that a pictographic Chinese basic shape might bear a visual resemblance to the meaning of the morpheme it graphematically represents is not of interest. Rather, pictography can also be understood in a semasiographic sense in which it surpasses language. This is evident from literacy acquisition: at an early stage, children do not understand that graphic shapes represent language but recognize them as resembling certain objects instead. Thus, they believe shapes directly represent visually similar referents. Accordingly, at the graphetic level, arbitrariness or non-pictoriality imply that basic shapes – regardless of the graphematic relations they take part in – do not directly resemble any external referents. Treiman & Kessler (2014: 171) note that children have difficulty acquiring arbitrary, abstract shapes since they cannot associate them with the shapes of any objects they are familiar with. The authors argue that some basic shapes, however, are known to children, such as |O| which children know as a circle, or |X| which some might know from playing tic-tac-toe. Such familiar shapes might be easier to learn and memorize than shapes that do not resemble anything. This observation is in line with findings of a large-scale study by Changizi et al. (2006). It shows that the basic shapes in the scripts of the world[235] exhibit topological configurations that are found in natural scenes. This, in turn, is in accord with Dehaene's (2009) by now well-known *neural recycling hypothesis*, which claims that a number of brain regions which were originally (and still are) responsible for other tasks were repurposed for reading and writing processes. For example, mechanisms of object recognition are exploited in the recognition of basic shapes. This explains why under human pressure, basic shapes evolve to resemble salient visual features of objects found in the environment of humans.

However, aside from incorporating line junctures reminiscent of natural scenes, basic shapes are predominantly abstract. The driving forces behind this are facilitation and the principle of economy: as abstract shapes do not have to resemble anything, they can be designed in a way that makes them easier to produce than more elaborate pictographic shapes. Notably, when it comes to the loss of pictography, ontogeny shares remarkable similarities with phylogeny, as in the course of their diachronic development, the world's scripts became increasingly abstract, gradually losing their pictographic character. Take Chinese: Figure 35 illustrates the increasing abstractness in the diachronic development of its basic

**234** It is necessary to specify that while writing is largely non-pictorial or non-pictographic, it is not entirely non-iconic as there are different types of iconicity. Pictography equals *imagic iconicity*, which writing, from a synchronic perspective, exhibits only to a small degree. By contrast, *diagrammatic iconicity* is not uncommon in writing systems (cf. Section 6.2).

**235** In one group, the study investigated 96 scripts which are used for non-logographic writing systems and, in another group, Chinese, nonlinguistic symbols, trademarks, and children's scribbles (cf. the criticism in Daniels 2018: 152).

shapes, resulting in a synchronic state that exhibits only little iconicity of this kind (cf. Xiao & Treiman 2012). This is reflected in ontogeny in the fact that when children are asked to *write* (vs. to *draw*), they usually do not produce recognizable drawings (cf. Gombert & Fayol 1992) – their scribbles are non-pictorial. Studies suggest that in literate communities, children learn to differentiate between writing and drawing before the age of 3, and as mentioned, their own attempts at writing are much less iconic than their drawings (cf. Treiman & Kessler 2014: 114; cf. Otake, Treiman & Yin 2017). For children, apparently, "pictoriality is perceived as a sufficient criterion for nonwriting status" (Lavine 1977: 92).

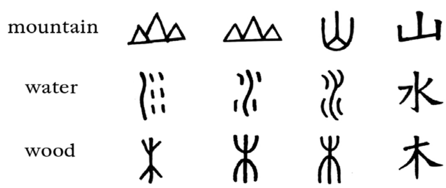



Figure 35. Increasing abstractness of Chinese script from left to right, adapted from https://buckinghamhsiao.wordpress.com/2018/12/05/drawn-words-pictographs-in-the-chinese-language-and-visual-culture-in-drawing-research-theory-practice-vol-3-2/ (September 23rd, 2020)

Treiman & Kessler (2014: 116) note that the visual makeup of writing distinguishes it from other visual systems, making it possible for children to differentiate between them. At a more advanced stage, the same skill – recognizing the systematicity and the systematic fit of a given script (cf. Section 5.1) – allows children not only to tell apart writing from non-writing but also to distinguish between different scripts as subcategories of writing. Thus, between the ages of 3 and 4, children identify as 'writing' basic shapes that exhibit the general features of the script they have been primarily exposed to (cf. Treiman & Kessler 2014: 168) and reject as writing basic shapes that lack these features. In her study, Lavine included three different classes of basic shapes: class I consisted of Roman script, class II of basic shapes sharing some features with Roman script (including, for example, basic shapes from Hebrew script), and class III included basic shapes from scripts such as Chinese or Maya which are to a remarkable degree visually dissimilar from Roman script. The results showed that even the youngest children (age 3) rejected basic shapes of class III as writing (cf. Lavine 1977: 93). The differentiation between the more similar basic shapes from class I and class II proved more difficult, and only 5-year-olds significantly preferred basic shapes from class I. Lavine (1977: 94) concludes that "it is the general features shared by members of the conventional set that are first picked up by children". Thus, it is not surprising that children accept the foreign Hebrew shapes as writing since they share salient features with the script they had been predominantly exposed to. Ganopole's study yielded similar results: 62% of 3-year-olds rejected geometric figures as writing. With increasing age, the ability to distinguish Roman basic shapes from numerals and geometric figures advanced, and 5-year-olds were able to differentiate correctly between shapes of all of those categories (cf. Ganopole 1987: 430). The studies of

Treiman et al. (2007) and Levy et al. (2006) are congruent with this, with the latter suggesting that 4;5 is the age at which children are able to distinguish at above chance level basic shapes of their script (Roman) from those of foreign scripts. A number of studies investigated the same skill of differentiation in Chinese children and found that they are also able to distinguish Chinese basic shapes from foreign basic shapes (from Roman script, Kannada script) or pictures at above chance level (cf. Qian et al. 2015; Zhang, Yin & Treiman 2016).

The features discussed so far – artificiality, arbitrariness, abstractness, visual blandness, and systematicity – are governed predominantly by the principle of economy (Treiman & Kessler 2014: 153–159) which is itself a byproduct of the forces of facilitation and, to some degree, homogenization.

#### Two-dimensionality, rectilinearity, and directionality

Spatiality is constitutive of writing, and two-dimensionality is one of its central characteristics. It is not primarily motivated by any of Watt's four forces of script development but instead by the inherent materiality of writing surfaces: as a product of graphic activity, writing extends in space, and it can do so reasonably only in two dimensions – one-dimensional writing would merely equal a straight horizontal or vertical stroke (and three-dimensional writing is hard to imagine; would it be like sculpting?). However, writing exploits the two dimensions not randomly but in a rectilinear fashion. This is grounded in cognitive constraints: if basic shapes were positioned on the writing surface in a chaotic manner, e.g. not along a line but randomly, this would greatly complicate writing and, even more so, reading. By comparison, if writing is laid out linearly, these processes are facilitated. Children learn about this fundamental feature of writing early on and tend to reject displays of writing in which units are not arranged along a line. In Ganopole's study, 87% of tested 3-year-olds rejected as writing strings of graphic material that were not arranged along a horizontal line. However, 3-year-olds could not explain their rejection of such displays, whereas older children offered explanations such as "It doesn't go the right way" or "It's got to go straight" (cf. Ganopole 1987: 426f.). Similarly, Lavine (1977: 92) found that linearly arranged shapes were judged as writing more frequently than nonlinear arrangements and concluded that children perceive linearity as a basic feature of writing. This finding was reproduced and further specified by Treiman et al. (2007: 1466f.) who showed that US children under the age of 4 rejected nonlinear arrangements of personal names written in uppercase shapes that were scattered randomly on the page. Among the linear arrangements, all children – including those who, in a prior reading task, had not been able to read any of the words presented to them – preferred the horizontal arrangement over other linear (e.g. vertical) arrangements. As for different linear arrangements, interestingly, the authors found that none of the children preferred vertical arrangements, which are prevalent in some writing systems (see below) over diagonal arrangements, which do not occur in any of the world's writing systems. They concluded that the so-called *differentiation hypothesis* (cf. Tolchinsky

2003), which is also known as the *universal-to-specific hypothesis* and holds that universal features of writing are acquired by children before system-specific features, must be inadequate (cf. also Treiman, Mulqueeny & Kessler 2015 and their findings). The hypothesis predicts vertical arrangements to be universally preferred over diagonal arrangements, which was, as mentioned, not the case.

Writing is not only linear but (synchronically) also oriented in a single direction. This *directionality* (or *vectoricality*, cf. Brekle 1994c: 136) is a graphetic feature crucial for a theory of writing. While the above-mentioned features of two-dimensionality and rectilinearity render writing linear, they prescribe no fixed direction, and indeed, in the writing systems of the world, multiple writing directions exist. Among horizontal writing systems, many are written from left to right (such as all alphabets using Roman script), while others are written from right to left (such as Arabic and Hebrew). Among vertical writing systems, some are written in vertical top-bottom columns from left to right (e.g. Mongolian), others from right to left (e.g. Japanese, Chinese). The specific direction of a given writing system appears to be acquired by children first in perception and then in production: in their study, Tolchinsky-Landsmann & Levin (1985: 329f.) found that the majority of 3-year-olds did not produce writing in a consistent direction. Interestingly, some children switched the direction of writing each time they reached a physical boundary of the writing surface and produced an arrangement known as *boustrophedon*, a directional principle that was used in several writing systems in the past (cf. Coulmas 1996a: 49; Figure 36). Among the 42 Hebrew children who participated in the study, children from the age of 4 had understood the unidirectionality of writing, i.e. wrote only in one direction. However, the conventional sinistrograde directionality of Hebrew was only grasped by children from the age of 5 (cf. Tolchinsky-Landsmann & Levin 1985: 336).

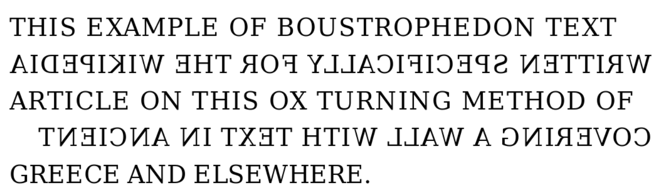



Figure 36. A writing strategy referred to as boustrophedon, from https://upload.wikimedia.org/wikipedia/commons/thumb/f/fc/Boustrophedon_%28all_caps%29.svg/1024px-Boustrophedon_%28all_caps%29.svg.png (March 24th, 2020)

With respect to vertical directionality, top-to-bottom appears to be a universal preference (cf. Treiman & Kessler 2014: 111), although, as Clay (1975: 13) notes, "when children were asked to 'Read it with your finger' they moved from left to right, right to left, top to bottom and bottom to top", implying that at some (early) point in literacy acquisition, bottom-to-top directionality is also considered at least an option. In fact, there exists no writing system that is written from bottom to top, which is likely due to the same reasons that children grasp the universal top-to-bottom-directionality at a younger age than the typologically variable horizontal directionality. Directionality affects a number of other features of writing, the most important of which is arguably the spatial and hierarchical organization and makeup of basic shapes, in other words: how their segments are arranged in space and how they are connected. This organization has been referred to as *topology*, and a specific subfeature is the *orientation* of basic shapes. While two-dimensionality and directionality are *relational features*, i.e. features that are relevant for a sequence of basic shapes, topology and orientation are *individual features* and can be assessed for individual basic shapes (cf. Section 1.3).

A revelatory example was mentioned in the context of the systematic fit of scripts: the reversal of basic shapes that do not exhibit the prototypical features of a script. Treiman & Kessler (2014: 173) suggest that effects such as the reversal of |J|,[236] the sole uppercase shape of Roman script that is neither vertically symmetrical (such as |M| or |A|) nor oriented rightwards, i.e. equipped with a coda on the right (such as |R| and |K|), can be explained by the *theory of statistical learning*. It holds that children observe and internalize the prototypical rightward orientation of uppercase shapes of Roman script, which leads to mistakes in production. This corresponds with Watt's view that |J| is an outlier in the Roman script. However, statistical learning might not be the (only) reason behind these reversals.

Clay (1975: 24f.; 64) notes that children, when given a surface to write on, are flexible in their choice of a starting point. This is what J.-P. Fischer (2017: 534) exploited in his study: French children were instructed to write their names as "beautifully as possible" and to start writing where an ink dot was positioned. In the "dot-left condition", this dot was located at the right end of the surface, which, due to the directional flexibility of the children, encouraged right-to-left writing. When writing from right to left, children reversed most of the right-oriented basic shapes such as |E|, resulting in shapes whose codas were oriented leftward. This calls for a reevaluation of the so-called *right-writing rule* (RWR, cf. Fischer & Koch 2016), which holds that children "in our left-to-right oriented culture [...] should prefer to orient their writing rightwards when writing the characters from memory". This rule, which is in line with the above-mentioned theory of statistical learning, was previously assumed to be the cause for reversals of |J| in conditions of default left-to-right writing. In light of J.-P. Fischer's (2017) results, however, the right-writing rule must be phrased more generally. What is key, it appears, is not that basic shapes are produced to exhibit the same orientation as the statistical majority of the script's shapes, but rather that the basic shapes' internal orientation aligns with the direction of writing – whichever that direction might be. J.-P. Fischer (2017: 538) calls this the *writing-direction-orienting rule*, which supersedes

236 Instances of these reversals can also be found in the diachronic development of scripts. An example is archaic Greek lambda which derived from the Phoenician basic shape |✓|. Since all other basic shapes of Greek script were right-oriented, the shape was reversed to |⌄|. Additionally, since most of the shapes had their codas at the top rather than at the bottom of the basic shape, lambda further changed to |⌃| (cf. Watt 1988: 204).

the *right-writing rule*. McIntosh et al. (2018: 680) claim that this "implies that the heuristic that children apply is not that most letters face rightward, but that most letters face in the direction of writing (and reading)". Notably, this interpretation (cf. the formulation 'most letters' in both parts of the sentence) is still compatible with the theory of statistical learning, as it merely argues that children are influenced by the fact that the majority of a script's shapes incidentally happen to face in the writing direction (which, in most cases, is rightward). Thus, the rule is not meant to explain why the shapes' orientation aligns with the writing direction and the alignment itself is also not regarded as the motivation behind the children's behavior. However, children's treatment of digits provides interesting evidence: even though, with respect to hand movements during production, the majority of digits is left-oriented (cf. |1|, |2|, |3|, |7|, |9| vs. the right-oriented |6|, |5|, |4|), in the default left-to-right writing condition, children more often reversed the left-oriented digits than the right-oriented ones. To explain why the digits with the statistically dominant feature [left-oriented] are reversed instead of the right-oriented ones, Treiman & Kessler (2014: 170) suggest that children, for some reason, treat digits and "letters" (scriptual basic shapes, cf. Section 1.2.1) as a single set in which [right-oriented] is the dominant feature (cf. also Watt 1983a: 381). In the light of J.-P. Fischer's (2017) results, however, I argue that it is more likely that the *writing-direction-orienting rule* actually reflects that an alignment between shapes' orientation and writing direction is well-suited for processing, and this overwrites the statistically dominant feature of [left-oriented] among the inventory of digits.

Additional evidence for the natural processing fit of the writing-direction orientation of basic shapes comes from diachrony (cf. Jeffery 1961: 43–50; Sirat 1988: 185–199). Brekle (1996: 483) underlines that the dextrograde writing of Roman script did not prevail because of an orthographic regulation of the writing direction; instead, he illustrates a trajectory of user-initiated changes that made the ancestral Greek script gradually easier to process. Thereby, he offers causal explanations that are precisely what a theory of writing should strive for. Brekle's (1996: 486) reconstruction starts with the assumption that the ductus of writing in Phoenician script was dextral, i.e. the motoric production programs of basic shapes were executed from left to right. By contrast, the direction of writing was sinistrograde, i.e. shapes were written in sequence from right to left. Thus, ductus and writing direction were incongruous as individual shapes were produced internally from left to right whereas sequences of basic shapes were written from right to left. This did not change when the Greeks adapted the Phoenician script around 800 BC.

At that time, there existed a scribal practice that Brekle refers to as *capovolto*[237] (cf. Figure 37). In this practice of writing, which resembles a snake's

237 *Capovolto* is Italian and means 'upside down'. Zinn (1950/51: 24) argues that this designation is not fitting as at that stage in the development of writing, one could not yet speak of an 'up' and 'down' on the writing surface. Instead, he refers to this practice as *parallel-geschlängeltes Schriftband* ('parallelly wiggled scroll'), or short: *Schlangenschrift* ('snake writing').

mode of locomotion, when scribes reached the writing surface's physical boundary, they did not break lines but first turned shapes by 90 degrees and then by 180 degrees, which means shapes were reversed not only horizontally but also vertically. This, of course, posed great cognitive challenges to writers. *Boustrophedon* (as illustrated in Figure 36) represented the next developmental step in which shapes came to "stand on their feet" (Brekle 1996: 487, my translation). The term comes from Greek βοῦς, *bous,* 'ox' and στροφή, *strophē,* 'turn' and implies "turning like oxen during plowing". This practice introduced the concept of line breaks, i.e. the succession of lines beneath one another. There exist a number of inscriptions from around 700 to 600 BC that are written in boustrophedon in which the direction of writing changed in every new line. Brekle assumes that scribes of that time interpreted the basic shapes as generally reversible along the vertical axis (i.e. exhibiting no fixed orientation).

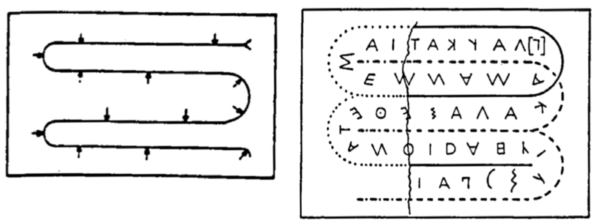


FIGURE 37. Capovolto writing (or snake writing), from Brekle (1996: 487–488)

However, *boustrophedon* was cognitively and physiologically unnatural precisely because it required changing the direction of reading and writing for every line. This cognitive unnaturalness led to optimizations regarding the above-mentioned naturalness parameter of writing-direction orientation. These changes can be reconstructed on the basis of Brekle's *hasta+coda-principle* (cf. Section 2.4). The visually dominant hastas (in most cases vertical strokes such as |l| in |b|) are accompanied by codas that face the writing and reading direction (such as |ɔ| in |b|). However, the fact that the codas visually (i.e. descriptively) face the writing direction does not necessarily reveal anything about their internal ductus, i.e. the way users have produced them, as the direction of writing entire basic shapes and the internal ductus of producing their constituent elementary forms might be incongruent (see above). This yields four theoretically possible combinations of the variables *writing direction* (sinistrograde vs. dextrograde) and *ductus direction* (sinistral vs. dextral): sinistrograde-dextral, sinistrograde-sinistral, dextrograde-dextral, and dextrograde-sinistral. The internal ductus of basic shapes, now, was predominantly dextral. This means that when writing in boustrophedon, scribes were required to change their motoric programs for every line: in sinistrograde lines, because of the dextral ductus, scribes first produced codas, followed by hastas (= sinistrograde-dextral), while in dextrograde lines, production started with the hasta and proceeded with the coda (= dextrograde-dextral). Eventually, from 600 to 500 BC and across Greece and Rome, the dextral ductus led to the gradual fixation of the dextrograde writing direction (cf. Brekle 1996: 489f.). This was arguably determined by the force of facilitation: fixing the dextrograde direction meant that all basic shapes could be produced with a ductus that corresponded with the writing direction, i.e. they could all be written dextrograde-dextrally.

In accordance with the evidence used in Naturalness Theory, Brekle (1996: 490) cites mistakes of writing found in inscriptions as support for the assumed universal preference of writing-direction orientation. Specifically, in some boustrophedon inscriptions, in sinistrograde lines, right-oriented shapes appeared although they should have only been produced in dextrograde lines.

What was discussed up until this point was individual basic shapes, sequences of basic shapes, lines, and the succession of lines. It is the latter that most demonstrably underscores the two-dimensionality of writing, which was largely neglected in the linguistic treatment of writing. However, as established in Chapter 1, writing makes use of both dimensions. At higher levels of organization, this concerns not only the succession of lines but, for example, also the arrangement of paragraphs or the arrangement of combinations of paragraphs of text and non-written material such as figures or pictures. These latter types of arrangement are generally referred to as *layout* (cf. also the concept of *graphetic dispositif* in Section 1.2.3). The arrangement of elements on a page is more challenging for children to grasp than the mere linearity of single strings of writing or even entire lines, and unsurprisingly, it is acquired later. Clay (1975: 39–41) observes that after children have successfully mastered the directionality of their respective writing system, they still run into problems concerning page arrangement. Figure 38 illustrates that a child may "[adopt] an easy solution and [fill] any left-over spaces with left-over words, ignoring at that moment any constraints of directional principles" (Clay 1975: 39). In this example, after writing 'we are going in the', there was no empty space left on the right side, so the child produced the final word 'bus' on the left side where there was still free space. Since the elements of a layout are arranged not simply on the basis of their materiality but usually based on their respective function in a text, the question of page arrangement is also relevant in the context of the linguistic fit, in particular for the parameter of indexicality (cf. Section 6.3).

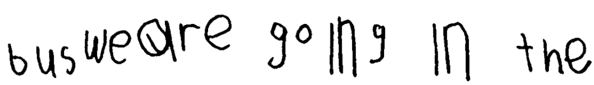



FIGURE 38. Problems with spatial arrangement, from Clay (1975: 39)

### Segmentality, finiteness, sequentiality and multiplicity, alternation

Units of writing are *segmental* and *discrete*. These features are based on the fact that writing is an analysis of language, and every writing system must choose which units of language it wants to represent in writing: phonological features, phonemes, syllables, or morphemes. Even when units of writing are connected graphetically, as is frequently the case in cursive handwriting or, in some writing systems such as Arabic, also in print, users are commonly aware of the segments that combine to form larger units of writing. This is arguably the most crucial difference between writing and speech, since in the acoustic continuum that is speech, no segments stand out (*a priori*).[238] Another universal feature of writing is its *finiteness*: every writing system and every script, for that matter, consists of a finite number of units. This, of course, is also a top-down reflection of the fact that the units of writing represent linguistic units, as the linguistic units that are represented by the basic units of writing come from (relatively) closed inventories such as the phoneme, syllable, or morpheme inventories of a language. By contrast, words or sentences, which are open classes, are never represented by basic units of writing (cf. Meletis accepted b). However, finiteness has its limits, and only abstract graphematic and graphetic units, i.e. graphemes and basic shapes, are finite, whereas graphs are not. When it comes to the concrete materialization of writing, actually realized visual shapes are subject to a degree of variation that is unprecedented in speech and that can lead to challenges in the modeling of categorical perception.

The segmental and discrete units of writing are produced in sequence. Accordingly, for something to be considered writing, multiple units need to co-occur, which is a feature that has been referred to as *multiplicity* (cf. Ganopole 1987: 419f; Lavine 1977: 93). In literacy acquisition, children are more likely to classify strings of basic shapes as writing than just single basic shapes: for example, they are more likely to accept |PLVN| as writing than just |P| (cf. Treiman & Kessler 2014: 111f.). This is in accord with Ganopole's (1987: 428) finding that across four age groups (English-speaking children aged 3, 4, 5, and 6, respectively), 90% of the children rejected displays of a single Roman basic shape as writing. Similarly, strings of two basic shapes were "troublesome" for many and were rejected "on the basis of insufficient quantity of symbols". Meanwhile, three basic shapes proved more acceptable, while strings of four were accepted by all children. By contrast, Lavine (1977: 92) showed that multiplicity was a significant criterion only in the youngest group of children, whereas it was not for 4- and 5-year-olds. She suggested as an explanation that "for the children as a whole, the unit itself takes on greater importance with age". This could reflect the fact that in many writing systems, there exist words that consist of only a single grapheme, such as <a> in English or monomorphemic words in Chinese. This knowledge is expected to be acquired later, at a stage at which children gradually grasp the relation between the visual and the linguistic.

Empty spaces between units constitute the segmentality of writing. While empty spaces are more or less visually salient at all graphetic levels of writing (cf. Chapter 1), the most salient is the empty space between 'words' – in those

238 At this point I would like to remind the reader that some scholars claim the phoneme is not an *a priori* cognitively real category and that its status as a segment of speech is epiphenomenal and the result of alphabetic (or, more accurately, segmental) writing systems (cf. Faber 1992; Davidson 2019).

writing systems in which written ‘words’ exist. Expectedly, this ‘interword’ space features prominently in studies on emergent literacy (e.g. Tolchinsky 2003: 69; 184f.; Tolchinsky-Landmann & Levin 1985: 324f.; Gombert & Fayol 1992: 38). Children in writing systems that exhibit interword spacing appear to perceive it as a salient feature of writing and employ empty spaces in their own early attempts at writing. What children have actually acquired at this stage, however, is merely a “surface property” of writing, i.e. a graphetic feature. Children learn about the graphematic function of interword spaces only at a later stage when they acquire knowledge about the representational function of writing as well as metalinguistic knowledge about the unit ‘word’.

Children appear to perceive early on that the script they have been predominantly exposed to in their literate culture is a more or less coherent system. This is reflected in the fact that they do not classify basic shapes that do not conform to the general features of their script as writing (see above). A pressing question in this context is when and how children understand that a script is a more or less[239] closed inventory. In this context, it is interesting to note that Clay (1975: 31f.) describes how children sometimes tend to make lists of all basic shapes they know. This behavior could be interpreted as an early awareness of the fact that there is a limited number of units in a script. However, children not only make lists of basic shapes but later on also of written words and sentences, which are, as mentioned, open classes. Clay calls this the *inventory principle*. I am not aware of any research that has dealt with the question of how children acquire knowledge specifically of the finiteness of scripts, but it is reasonable to believe that it is acquired as children understand that basic shapes are employed graphematically, i.e. as visual signantia of graphemes. Only top-down, then, do they realize that since there is a limited number of sounds (or morphemes, etc.) that are represented by graphemes, and thus, a limited number of graphemes, the inventory of basic shapes must also be finite. As implied above, at an abstract level, scripts and grapheme inventories are finite, whereas, at the concrete graph level, variation is almost infinite. It is an obvious assumption that children grapple with this kind of visual variation and the challenges of categorization that it entails (cf. Treiman & Kessler 2014: 181). In production, children tend to test “how far they can go before a symbol changes so significantly that the meaning is altered” (Kenner 2004: 87). In other words, they experiment with the scope and boundaries of the graphetic solution space (cf. Section 1.2.1; see also below).

A given unit of writing is commonly not frequently repeated in immediate sequence, meaning it is much more likely that a unit is followed by a different unit rather than being repeated. There are exceptions to this, of course, such as frequent doublets in Finnish or only recently orthographically licensed German spellings which include the same grapheme produced three times in sequence such

239 “More or less” refers to the fact that new shapes can theoretically be added to scripts, which, however, is rather seldom the case at an advanced stage of development when a script has been in use for a long time.

as in <Schifffahrt> 'boat trip'. There is evidence that children acquire knowledge about the prototypical lack of internal repetition, which could be referred to as *variety* or *alternation*, even before being aware of the graphematic relations between basic shapes and linguistic units: even at this preliterate stage, children are more likely to accept strings of basic shapes that exhibit variation, e.g. ABC, than strings in which basic shapes are repeated, such as AAA (cf. Ganopole 1987: 428f.; Lavine 1977: 92). Children thus appear to have understood that "adjacent symbols within a string don't normally have the same shape" (Treiman & Kessler 2014: 112). Interestingly, Clay's account, which is, however, qualitative and observational in nature and includes neither concise information about data collection nor any statistics, mentions a so-called *recurring principle*. It holds that a child "who knows only a few letters or words can take a short cut to a long statement by repeating the same symbol again and again and again" (Clay 1975: 21). Notably, this does not necessarily contradict the above-mentioned findings. A possible explanation is that production lags behind perception and children are able to distinguish a greater number of basic shapes in perception than they do in production. The chronology of acquisition could be the following: children first acquire knowledge about the feature of multiplicity, while knowledge of the lack of internal repetition comes only at a later stage. Repetition of units at an early stage might merely be a reflection of the fact that children want to produce longer strings of writing even when they are not yet familiar with many units, and "as a self-initiated process[,] repetition can provide a wonderful sense of accomplishment" (Clay 1975: 21). However, counterevidence comes from newer, empirically grounded findings from Pollo, Kessler & Treiman (2009) who show that when children randomly produce and combine basic shapes they had acquired, they produce double sequences of the same basic shape less often than expected.

In an interesting study regarding repetition, Lehtonen & Bryant (2005) tested 121 Finnish children between the ages 6;5 and 9;8. The study's most relevant finding is that "already at the beginning of the first school year children possessed formal knowledge of doublet use and knew that word-initial doublets are not allowed" (Lehtonen & Bryant 2005: 211). This is striking as it suggests that even children who have not (yet) acquired the representational function of writing are already aware of positional constraints. Interestingly, doublets (such as |aa|) were not generally rejected – only the ones in positions that are not graphotactically licensed in Finnish. Children had apparently acquired a graphematically graphotactic constraint solely through exposition to the graphetics of writing, before (or on the verge of) understanding that this constraint is indeed determined graphematically. Fittingly, Clay (1975: 22) underlines that when children are older, they must learn about specific constraints of English such as that no more than two graphemes are ever repeated in sequence. This, however, represents a genuine system-specific feature of the English writing system rather than a universal graphetic feature of writing.

Let us now turn to categories of writing that are more specifically linked to production before turning to those relevant in perception.

## 7.1.2 Production

One method of organizing the analysis of graphetic categories relevant for a theory of writing is to investigate them with respect to the central conflict between the needs of the writer on the one hand and the needs of the reader on the other. This conflict is associated with Watt's three forces of homogenization, facilitation, and heterogenization, and by Treiman & Kessler's principle of economy (see above). Homogenization and economy remain vague as to the actors that drive them. Even though it might appear obvious that they are driven purely by production, they actually refer to both production and perception. For example, the principle of economy can, on the one hand, pertain to the number of motoric programs one needs to memorize in order to produce the basic shapes of a script or, on the other hand, to the number of visual templates (or features) that one needs to memorize and identify to recognize them. Homogenization, ultimately a cognitive force, can, as already mentioned above, also be seen from a perceptual point of view.[240] A central hypothesis regarding the systematic fit of scripts holds that scripts with visually homogeneous (but, crucially, sufficiently distinctive) basic shapes are processed by users with less effort than scripts with more heterogeneous shapes.[241] Heterogenization, on the other hand, appears to be an exclusively perceptual force, as it ensures that the shapes of a script do not grow too visually similar to be distinguished as separate units. In case they did become too similar, a secondary process would be necessary in which they are made more distinctive again, which occurred in the history of Arabic script (see below). Finally, inertia or, in the list of Treiman & Kessler's (2014: 159) principles, conservatism, as the fourth and final force, is not a driving force of change but instead the inhibitor constraining it. As such, inertia serves as a crucial force assuring the stability of scripts and writing systems. This results in the fact that writing systems change much more slowly than the language systems they represent, which can lead to the increase of opacity in the relations between the two – in other words, a decrease of the linguistic fit.

In this section, I will discuss which categories and processes involved in production affect the processing fit (cf. also Sirat 1988 for the 'physiology of writing'). In the next section, I will do the same for perception. It must be noted that productional and perceptual aspects can rarely be separated completely, and crucial interactions between them will have to be dealt with throughout. What can

240 For Watt (1983b: 1545), homogenization is indeed mostly perceptual, for it "operates mostly on the cognitive phanemic grammar, altering the remembered forms of letters (or, equivalently, the rules for determining the forms of letters)".

241 For the concrete level of graphs, there is indeed evidence supporting this hypothesis. In a number of studies, Sanocki (1988, 1991; Sanocki & Dyson 2012) found that visual uniformity in the design of individual typefaces – achieved, for example, by means of stroke weight, stroke contrast, and stress angles in the design of graphs materializing different basic shapes – exerts a beneficial effect on processing. So-called *font regularity* (Gauthier et al. 2006: 555) helps the processing system "[tune] itself to exploit regularities of a font" (Sanocki & Dyson 2012: 133). This 'skill' on behalf of readers is referred to as *font tuning* (Gauthier et al. 2006: 541).

be stated in advance is that facilitation is the driving force behind making production more suitable to human processing needs. Notably, the following paragraphs are focused exclusively on processes involved in handwriting. However, handwriting, today regarded by some as a "forgotten language skill" (Medwell & Wray 2008), is, in many literate cultures, used much more rarely than before as it has in many contexts been replaced by electronic, i.e. digital ways of writing. These ways of writing include typing and swiping and must be included in a theory of writing. They will be addressed in the next step.

A central consideration for the facilitation of handwriting is the motoric complexity of basic shapes. It is affected, for example, by the number of elementary forms in a basic shape as well as the nature of connections between them. However, segmentation of basic shapes is only one part of production, as the order of production of elementary forms is also central (Thomassen & Tibosch 1991: 269). It is obvious that in general, "writers want shapes they can produce quickly, with little cognitive and muscular effort" (cf. Treiman & Kessler 2014: 153; cf. also Thomassen & Tibosch 1991: 270). 'Little effort' is most often associated with a reduction of elementary forms that make up basic shapes. From a diachronic perspective, this reduction – the productional result of economy – often promotes the homogenization of a script's basic shapes. Here, it is paramount to distinguish between the homogenization of motor programs involved in the production of basic shapes, which is the result of reduction, and the homogenization of the visual appearance of basic shapes, which can lead to a decrease of visual distinctiveness and, thus, perceptual ambiguity. In fact, productional facilitation merely dictates that the basic shapes become easier to *write*, which is why the homogenization of motor programs across basic shapes does not necessarily equal homogenization in the corresponding visual shapes – similar movements need not produce similar shapes. On the contrary, facilitation might even render shapes visually more distinct while they become easier to produce, although this appears to be constrained by the perceptual aspect of the force of homogenization. In any case, the interaction between productional and perceptual forces is complex.

Aside from the production of individual basic shapes – which is rather the marginal case in writing processes – and following the conception of natural processes (cf. Section 4.2.2), the production of sequences of basic shapes is central to the processing fit as there appear to exist effects of coarticulation in writing (see below).

A model of the processes involved in handwriting that has remained influential is Goodnow & Levine's (1973) *grammar of action*. In it, the sequence and characteristics of children's copying processes are described, and the motoric processes of production captured by it are said to have a "huge constraining power" (Thomassen & Tibosch 1991: 278). They can be summarized as follows:

| | |
|---|---|
| a | Start at the top of the pattern |
| | Start at the left-hand extreme of the pattern |
| | Start at a vertical segment |
| b | Draw strokes downwards |
| | Draw strokes rightwards |
| c | Thread: continue pen-down |
| d | Anchor: connect to earlier strokes |
| e | Draw parallels in immediate succession |

(Thomassen & Tibosch 1991: 271)

One group of processes is concerned with the *starting point* (a), another group captures *stroke directions* (b), and the final group is dedicated to *stroke sequences* (c-e) (cf. Thomassen, Meulenbroek & Hoofs 1992: 72). Stroke sequences include the central production strategies of *threading*, i.e. continuing, when possible, to write with the pen touching the surface and thereby avoiding to lift the pen, and *anchoring*, i.e. using an earlier, already produced stroke as a connecting point for a successive, connecting stroke and thereby avoiding to start a new stroke in empty space. It is important to note that not all of these processes are universal: "start at the left-hand extreme of the pattern" and "draw strokes rightwards" only apply to right-handed individuals in writing systems with a dextrograde writing direction. The others, however, are expected to apply universally and have been described for the production of the basic shapes of Roman script (cf. Parkinson, Dyson & Khurana 2010) and Chinese script (cf. Flores D'Arcaïs 1994). These dynamic processes of handwriting imply that *top-bottom directionality*, *verticality*, *continuity*, *connectivity*, and *repetition* are relevant categories of the productional processing fit. Some universally preferred visual features of basic shapes might be explained by them: the predominance of cardinal strokes as opposed to oblique strokes could be a result of the preference for a downward direction of stroke production, and the preference for connected segments of basic shapes as opposed to unconnected ones could be a result of anchoring. These features will be discussed below with respect to their relevance for perception. Note that some of these above-mentioned processes are in conflict with each other: in the production of |E|, for example, the same movement has to be repeated three times (and the parallel production of the three horizontal strokes is economic, per rule (e)), but this repeated movement requires lifting the writing instrument each time, violating the process of threading. Thus, even among basic processes of handwritten production, conflicts can be detected.

Research on factors that cause difficulties in handwritten production is rather scarce. Accordingly, a recent study by Gosse et al. (2018) promises to "[offer] for the first time a universal classification of the graphic characteristics of words" by which it claims to "[enable] the quantification of the graphic complexity of words". The results of the study, in which twenty French second-grade children participated, show that the following factors – listed in order of impact – influenced the legibility of children's handwriting negatively: (1) modified links, (2) angles, (3) curves, (4) pen-ups, and (5) length. The only of these factors that had a negative effect on handwriting speed was angles.

An example of a (1) modified link is the connection between |b| and |r| in connected cursive handwriting (cf. Figure 39). In the study, such modified links caused aesthetic distortions, incorrect size, and incorrect relative height of shapes as well as poor alignment (cf. Gosse et al. 2018: 1201). The authors conclude that some sequences of basic shapes, when realized as concrete graphs in handwriting, require special links that change the shape of individual graphs (which is analogous to ligatures in typography). From a productional perspective, it is surprising that modified links would be difficult for children, as suggested by distortions in the children's production. What would rather be expected is that the process of connecting graphs in a way that enhances handwriting fluency and does not require lifting the pen is better suited for processing than materializing basic shapes in a sequence exactly the same way they would be produced in isolation. Imagine that no positional allographs existed in Arabic: regardless of their position within a word, the graphemes would always be represented by the same basic shape. Such a situation would likely be easier to process cognitively, as fewer basic shapes would need to be remembered. However, from a productional perspective, it is the connectivity of the positional variants that allows fluency in handwriting – a benefit that would be lost. In the light of all this, the existence of the positional allographs in Arabic script furthers the impression that its development was, for a formative period of time, shaped by the needs of scribes rather than those of readers. This apparent primacy of production even led to the visual collapse of several basic shapes that in a subsequent step needed to be disambiguated by dots (see below).

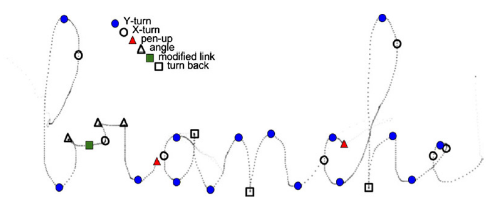



Figure 39. Example of a modified link between |b| and |r|, from Gosse et al. (2018: 1191)

The fact that in Gosse et al.'s (2018) study, modified links appear to have a deleterious effect on handwriting points to the possibility that written production is fundamentally constrained by perception. To explain this, the authors state that "children are taught at school to handwrite the letters of the alphabet in an invariant way, with the shape of the letters expected to be unchanging" (Gosse et al. 2018: 1201). Thus, children acquire one visual template per basic shape. At this stage of the children's development (in second grade), the pressure to keep this invariant visual template intact in their own handwriting appears to be stronger than the articulatory drive to facilitate hand movements (cf. also Morin 2018: 674). For Watt (1988: 201; 205), too "the program is merely a servant to the pattern", in other words: "the kinemic [= motoric, D.M.] side of the alphabet exists only to serve (bring into accessible existence) the phanemic [= visual, D.M.] side". In accord with this, Changizi et al. (2006: E120) claim that "human visual signs are selected for vision at the expense of motor". Thus, the results of Gosse et al.'s (2018) study, along with additional evidence discussed below, suggest that in the processing of writing, perception is primary (cf. also Primus 2006: 10; Watt 1988).

Since the complexity of the cognitively stored visual template affects production, different scripts require varying levels of productional competence. Children acquiring Chinese script, for example, have to show more pen control than children acquiring Roman script as they "also need to be able to recognise small differences in stroke patterns, to check that they have written each character correctly" (Kenner 2004: 76). The graphetic solution spaces for basic shapes of the Chinese script are much more restricted, meaning that "a small difference in the stroke pattern can make it look like another character with a different meaning" (Kenner 2004: 76).

The second variable in Gosse et al.'s (2018: 1202) study that caused children problems in handwriting was angles, which take children longer to master. The authors suggest that this difficulty could be caused by the fact that "changes of pen-stroke trajectory induced by the angles could be more demanding" graphomotorically. Angles ultimately represent ruptures in motor programs. They do not necessarily have to be *lifts* as not all angles require lifts of the pen and thereby violate the process of threading; however, if they are not lifts, angles represent *stops* (cf. Paz-Villagrán, Danna & Velay 2014). Stops are defined as "any discontinuity in graphomotor activity without the pencil having been lifted off the page and excluding any unavoidable inter-letter penlift" (Bonneton-Botté et al. 2018: 929). 80–90% of stops occur at angles, while stops within a single straight stroke are rather rare. Accordingly, curves, which are continuous, are expectedly easier to produce than angles. In fact, in Gosse et al.'s (2018) study, curves on the x-axis (such as in |u|), were even found to facilitate handwriting. This finding is interesting insofar as *roundness* vs. *angularity* were identified as possible base categories for a script typology (cf. Section 1.3). As these empirical findings suggest that curves are easier to produce, one could hypothesize that scripts consisting of predominantly round basic shapes (such as Georgian or Telugu) might be easier to write (i.e. have a better productional processing fit) than more angular scripts such as Korean Hangul. It is important to underline that Gosse et al. (2018) only evaluated Roman script, which features both curves and straight lines in its basic shapes and cannot be evaluated as either predominantly curved or angular. Thus, studies that compare handwriting processes in predominantly curved vs. predominantly angular scripts are needed to further investigate this hypothesis. The perception of curves, which might tell a different story, will be discussed below.

That curves exhibit a good productional processing fit is reflected not only in ontogeny and in performance but also in the diachronic development of (some) scripts. In this context, Watt (1994b) describes the process of *curvilinearization*, the "rounding off" of angles (cf. Watt 1980: 13). His main claim is that the development of curves from what were originally angles is due to facilitation of production processes (cf. Watt 1980: 13f.). Notably, there is also the opposite development, so-called *angularization*, which raises interesting questions. Runes are a pertinent example, as they angularized the curves of Roman script (cf. Treiman & Kessler 2014: 158). As they were produced on wood, one hypothesis holds that paragraphetic factors – the writing surface, the tools used, and the mode of production itself – led to the production of more angular forms. Note that, if this were indeed the reason, it would represent an external explanation for angularization that does not contradict the fact that curves are generally easier to produce. An interesting implication of this that will be addressed below is that a particular paragraphetic production scenario – e.g. carving on wood – can be less suited to production than a different one that allows the production of curves – e.g. writing with a pen on paper. The use of specific materials in a given situation might of course be imposed by other factors – such as, simply put, availability of materials. These factors are treated in the context of the graphetic sociocultural fit (cf. Section 8.1). The crucial point is that these sociocultural factors have the power to overwrite the processing fit. Another example of angularization that might not be as easy to explain is the fact that Chinese script, as it grew more abstract over time (cf. Figure 35 above), also lost much of its curvilinearity and became increasingly angular.

One of the graphetic features of writing identified as problematic for handwriting in the previous section is its segmentality. As implied by the preference for *threading* in the *grammar of action* and proven by studies, during the acquisition of handwriting, children experience difficulties with discontinuity, i.e. lifts of the writing instrument (cf. Treiman & Kessler 2014: 154; Gosse et al. 2018: 1202). Children's first writing attempts are sometimes characterized by unbroken wavy lines that do not allow a visual segmentation in separate units. Treiman & Kessler (2014: 111) suggest that this does not necessarily imply that children have not acquired knowledge of the segmentality of writing but that they instead simply avoid lifting the writing instrument. At a global level, the articulatory preference for continuity has consequences for the visual makeup of scripts, one of which is that the elementary forms making up basic shapes are usually connected. Accordingly, basic shapes in which segments are unconnected, such as |Ξ| in Greek script, are rarer in the world's scripts. And crucially, if such unconnected shapes do exist in a script, they are often produced in a continuous (connected) way in handwriting (cf. Wang 1958: xxv-xxvi).

Chinese is an interesting case with respect to the sequence of movements in handwriting, as in Chinese script, the stroke order is orthographically regulated. While sequential errors in the production process might not be visible in the final product,[242] they have other repercussions. This foreshadows one of the more specific questions asked within the broader scope of the processing fit of

242 Sometimes, sequential errors may be visible. In one of the Chinese classes I took, when we were trying out calligraphy with brushes and ink, I could not hide that I had neglected to follow the correct stroke order. My teacher had not witnessed the actual production of the character I had attempted, but in the product, the nature of the brush strokes clearly revealed to her that I had produced them in a wrong sequence (and direction, on top of that). A wrong production sequence can have detrimental effects on recognition, as Coulmas (1996a: 480) notes: "[...] in the absence of a fixed

orthographies (cf. Section 7.3): Can orthographic regulations negatively influence natural practices (in this case a natural sequence in the handwritten production of basic shapes)?

Facilitation of processes involved in handwriting leads to the emergence of informal variants of scripts. These are "flowing, quick, and nonmonumental" (Treiman & Kessler 2014: 155) and are often referred to as *cursive*. As a diachronic process, cursivization is a "well-known evolutionary [trend]" (Morin 2018: 674). Examples include the development of demotic script out of hieroglyphs (cf. Coulmas 1996a: 124) as well as the so-called "minusculization" of Roman script, i.e. the systematic development of lowercase basic shapes out of uppercase basic shapes that was led by the acceleration of production processes and a switch of writing surfaces from stone to papyrus and parchment (cf. Wiebelt 2003: 301). In the course of this development, the mentioned *hasta+coda-principle* played a central role (cf. Mallon 1952; Brekle 1994a, 1994b, 1995, 1998). Concerning minusculization, it is not only the dynamic processes involved in the production of lowercase basic shapes that are more natural (in that they are more economic, cf. Treiman & Kessler 2014: 180) than the processes involved in the production of uppercase basic shapes but also the static product that is more natural in some respects. Accordingly, as Simpson et al. (2013) note, uppercase basic shapes exhibit less intra-inventory similarity than lowercase letters, meaning that their systematic fit is not as natural. In line with this, some scholars argue that lowercase basic shapes are the default variants while uppercase basic shapes are marked (cf. Brekle 1998: 1; Primus 2004: 243, 2006: 9), although their diachronic development is actually inverse. Whereas in the case of the minusculization of Roman script, a whole new inventory developed out of the initial script and took on distinctive functions, sometimes, cursive variants continue to coexist as informal variants of formal basic shapes. In Chinese, for instance, among calligraphic styles, there are semi-cursive and cursive script styles in which the basic shapes appear drastically different from the shapes they take on in other, non-cursive variants of the script. In these cursive styles, the elementary forms of individual graphs are more connected than they are usually, and in some cases, the entire graphs themselves are also joined with each other (cf. Treiman & Kessler 2014: 155f.; Wang 1958; cf. Figure 40). Other examples of cursivization include ligatures: they develop out of two units that are intricately linked to such a degree that they gradually become to be perceived as one unit. That way, Latin *et* became the ampersand symbol |&| (cf. Tschichold 1981) and the German sequence of |ʃʒ| (or |ʃs|) merged and became the basic shape |ß| (cf. Häffner 2013). Notably, ligaturing is also a very common feature in many scripts employed in abugidas (cf. Share & Daniels 2016: 25).

order of strokes, handwriting would be impossible to read" since the sequence also assures that the product stays visually stable.

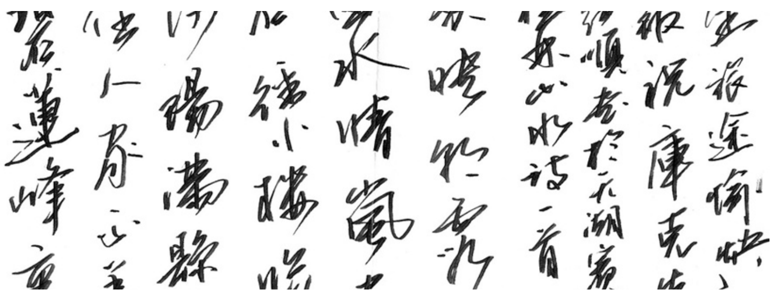


Figure 40. Chinese cursive script, from https://blog.tutorming.com/hs-fs/hubfs/learn_chinese-copy-copy.jpg (September 23rd, 2020)

A factor that interacts with facilitation in important ways is frequency. Research suggests basic shapes that occur more frequently in the use of a writing system tend to be simpler and relatively smaller in size than less commonly used basic shapes. For example, frequently used basic shapes in Chinese have a smaller number of strokes than basic shapes that are used rarely (cf. Shu et al. 2003).

In a nutshell, for handwritten production, from the perspective of the product, what is most natural is a small number of curved, continuous segments that should ideally be connected. As for the orientation of basic shapes, if they exhibit codas, these should most naturally be located in the direction of writing (or should allow production in the direction in which users typically write).

Turning to typing processes, it must first be noted (rather trivially) that they differ from handwriting processes in significant ways. This is true despite the fact that technically, typing is also "writing by hand" since it also involves the hands as articulators. However, nothing *per se* is graphically produced by the hands themselves when the fingers press the keys of a keyboard or typewriter.[243] Crucially, thus, optimizing or facilitating typing processes does not involve changing the shapes. In this vein, Treiman & Kessler (2014: 177) note that "typists can't get faster by modifying the shapes of letters, but they can press the keys more quickly". This implies that becoming faster in pressing the (correct) keys is an analog to writing faster in handwriting. Beginning typists are slower and work sequentially, as they usually attempt to locate the keys in succession. Usually, at this initial stage, finding the right successive key involves looking at the keys (cf. Treiman & Kessler 2014: 177). Later, when typers have constructed a "cognitive map of the keyboard" (Longcamp, Zerbato-Poudou & Velay 2005: 77), the location of keys can be retrieved from memory. When it comes to the coordination of movements the hands must make, the naturalness of the processes involved varies depending

243 Technically, also in handwriting (e.g. with a pen), nothing is produced by the hands themselves. It is rather the pen – led by the hands, of course – that produces the graphic traces. However, the crucial difference between handwriting and typing is that in the former, the mediation of the hands is much more direct than it is in typing since in handwriting, it is the actual movements of the writing hand that constitute the form of the product, which is not the case in typing.

on the distance between keys and the associated switching of hands that are used to press them. Typing a sequence of keys is most natural when (1) the next key to be pressed is pressed by the opposite hand, less natural when (2) a different finger from the same hand is used and least natural when (3) the same finger (from the same hand) is used. The varying degrees of effort involved in these three different scenarios coincide with the time needed for executing them (cf. Treiman & Kessler 2014: 177; Gentner 1983; Terzuolo & Viviani 1980). The fact that writing instruments generally influence production processes to a remarkable degree and that in typing, the distance between keys is crucial, renders the layout of keyboards a central factor in the evaluation of the naturalness of typing. The most widespread keyboard layout for Roman script is the QWERTY layout. It was "designed in 1873 to minimize jamming of the keys by maximizing the distance between frequently typed pairs of keys". However, it was still designed "without regard for ease of learning or typing" (cf. Rumelhart & Norman 1982: 2). Since individually, all keys are the same (and, if location is disregarded, the pressing of all keys is the same), it is not easier to produce an |x| than it is to produce an |o|. Arguably, thus, natural processes in typing can be evaluated only for sequences of key presses. Here, what is paramount, as mentioned, is the design of the instrument, the keyboard (cf. also Noyes 1983).

At first glance, due to their drastically different nature, it might appear inadequate or impossible to compare processes involved in handwriting and typing. However, in one respect, they are comparable: the effect they have on the memorization of shapes. More generally, this concerns the influence that either type of production has on cognition and, consequently, perception. This is a question of broad relevance. In fact, it has even attracted the public's attention in that it is being utilized in discussions centering around whether handwriting or typing has more cognitive benefits. In this context, three findings are particularly worth mentioning: The first comes from a study by Wong et al. (2018) that focuses on one of the inter-inventory free allographs of the grapheme <g> (cf. Section 2.3), so-called *looptail g* (or *closed-loop g*), which takes the form |g| as opposed to *opentail g* (or *open-loop g*), which appears as |g|. Many participants of the study were not even aware that a second variant that differs visually from opentail |g| existed when asked whether uppercase |G| has two lowercase equivalents. Indeed, open-loop |g| is common in a great number of typefaces (predominantly sans-serif) and, more importantly, in handwriting. Even after participants were instructed to actively search for instances of the grapheme <g> in a text with multiple instances of looptail |g|, most of them could only produce opentail |g| in a subsequent production task; only one participant correctly produced a looptail |g|. Most strikingly, in a final experiment, when participants were presented with four possible choices of what looptail |g| could look like, one of which was correct while the other three served as distractors (cf. Figure 41), participants performed poorly in recognizing looptail |g|. What is striking about the findings of this study is that they suggest looptail |g| is, as ScienceDaily sums it up, "a letter we've seen millions of times, yet can't write".[244] In fact, looptail g is more common in printed materials than opentail |ɡ| (cf. Wong et al. 2018: 1331). Yet, despite this "massive visual experience" (Wong et al. 2018: 1332), participants failed in the described tasks. One of the possible reasons suggested by the authors is a lack of production experience with respect to looptail |g|. Accordingly, the findings invite the assumption that producing a basic shape contributes to *letter-shape awareness* (which I will call *basic shape awareness* in the following). The authors' model includes an *abstract amodal representation of a basic shape's identity* (which they formalize as {G}), *stored allograph representations* (in this case |g| and |ɡ|), and a *graphic motor plan* (cf. Wong 2018: 1328). Producing an allograph as a concrete realization of either |g| or |ɡ| helps strengthen the associated allograph representation. This, however, in the case of looptail |g|, fails to happen on the basis that users produce it very rarely. A conclusion associated with this is that "letter-shape awareness does not always accompany the ability to recognize a letter shape via automatized reading processes" (Wong 2018: 1331). All of this implies that handwriting could be necessary or at least beneficial for stable cognitive and visual representations of basic shapes and allographs and that there is a "tight coupling between the visual and the sensorimotor perception of letter shapes" (Longcamp, Zerbato-Poudou & Velay 2005: 68; cf. also Zemlock, Vinci-Booher & James 2018).

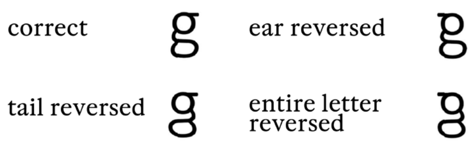



FIGURE 41. Correct |g| and distractors used in the experiment, from Wong et al. (2018: 1330)

The second example pertinent to this question concerns a fairly modern phenomenon termed *character amnesia* (cf. Xu 2015). It mainly 'happens to' users of the Chinese and Japanese writing systems who prefer to use modern, phonetic IME (Input Method Editors), methods that allow writers to input Chinese or Japanese graphemes by means of a phonetic Romanization (i.e. alphabetically). For example, when Mandarin users type "ma" in Roman script, they are presented Chinese graphemes with the pronunciation /ma/ (and one of the tones), usually in order of their frequency of use. From these graphemes, they have to choose the one they intended to write. It is argued that the widespread use of such input methods has side effects, one of which is character amnesia, defined as

> a state of affairs in which speakers of Chinese and Japanese, who are fully capable of *writing* (and obviously reading) almost any word they want by electronic means, often find themselves unable to *handwrite* correctly many of the same words if and when the need arises. In other words, they might experience this as a state of 'it's on the tip of my tongue' or, more accurately, 'on the tip of my pen'.
> (Almog 2018: 2, emphasis in original)

244 Cf. https://www.sciencedaily.com/releases/2018/04/180403140403.htm (May 6th, 2020).

The reason for this is that users of phonetic input systems seldom handwrite in the 'traditional' sense, i.e. very rarely or almost never use the graphomotoric programs for the basic shapes of Chinese and Japanese script. Consequently, in accordance with the provocative statement "use them or lose them", they forget them. They are still capable of recognizing these basic shapes and the graphemes they embody, which is necessary for using phonetic input systems, but can no longer write them by hand (which is a similar but of course not equivalent situation to the one described for looptail |g| above). Relying on alphabetic electronic input methods and a failure to (occasionally) produce the shapes by hand apparently leads to an impoverishment of graphomotoric programs. Dealing with writing exclusively through reading and typing could thus – and this is probably an exaggerated dystopian claim – gradually lead to forgetting how to write by hand entirely, making handwriting truly a "forgotten language skill" (Medwell & Wray 2008).

Two further studies are noteworthy in the context of comparing handwriting and typing. In the first (by Longcamp, Zerbato-Poudou & Velay 2005), two groups of children aged 3-5 were tested: one of them copied Roman basic shapes by hand (i.e. was handwriting), the other one by typing. The goal was to investigate whether movements involved in handwriting, as suggested above, contribute to a stable mental representation of basic shapes that serves processes involved in the recognition of these shapes. Results showed that in older children who had had more experience in handwriting, handwriting indeed led to better results in subsequent basic shape recognition than typing. In the second study, Frangou et al. (2018) instructed Finnish students to transcribe dictated stories using either (1) a pencil, (2) a keyboard, or (3) a virtual touchscreen keyboard. Thirty minutes after the task as well as one week later, the students' recollection for each writing task was evaluated, leading to the finding that handwriting resulted in a significantly better recollection both thirty minutes and one week after the initial dictation. This is interpreted as further evidence that handwriting has benefits for long-term memory. It is arguments such as these that are central in the public discourse on handwriting and the dreaded "demise" of handwriting (for discourses on handwriting, cf. Gredig 2020). Handwriting is claimed to have more cognitive benefits than typing, and this is often one of the cited reasons that it needs to be "saved". However, the practice of reducing handwriting to its cognitive benefits is also criticized by some who argue that handwriting, as a cultural technique, is also valuable in and of itself.

To conclude this section on graphetic production, I want to briefly mention one of the most modern forms of writing: *swiping*. Touchscreens – on smartphones, tablets, etc. – not only allow but even invite a so-called *direct touch* (cf. Ruf 2014: 51). Usually, there is no pen,[245] keyboard, or other writing instrument between the finger and the writing surface, which in this case is the screen (cf. Mangen 2016). The fingers directly touch and move on the screen, which has

245 There exist pens that can be used on touchscreens. However, these pens are less often used than fingers.

led to a gradual development of a grammar of so-called *touch gestures*, conventionalized movements with specific functions. Actual *handwriting*, i.e. producing graphomotor programs on the screen, is much less frequent in the context of swiping. Instead, text is entered through virtual keyboards that are displayed on the screen. In any case, swiping has introduced users to a new type of surface feel and requires specific fine motor skills. Given its recency, research on swiping is unsurprisingly scarce; however, as more is discovered about it, it must imperatively be integrated into a theory of writing. This concerns especially swiping's effects on motor skills and other modes of production of writing (for a review, cf. Wollscheid, Sjaastad & Tømte 2016). Some studies suggest, for example, that the smoother the surface of digital devices is, the harder it is to write on it because it provides lower friction (cf. Gerth et al. 2016a, 2016b). Furthermore, from a receptive or perceptual perspective, screens as new (writing and reading) surfaces should not be discarded as inherently less suited for reading than paper, which also appears to be a strand of public discourse. Newer studies have, in fact, suggested that when text display conditions are well matched between paper and screens (in the case of the cited study, these were tablet screens), there is no "reliable difference in reading between the two media" (Hermena et al. 2017: 1).

In general, this discussion highlights the importance of a paragraphetic perspective (cf. Section 1.2.4) and an investigation of writing materials. Materials and practices of 'inscription' have recently become the subject of a collaborative research center entitled *Materiale Textkulturen* (transl. 'Material text cultures') focusing on the materiality of writing in non-typographic societies. The eponymous book series published by De Gruyter has spawned thirty books so far.[246] The inaugural volume (cf. Meier, Ott & Sauer 2015) includes parts on materials and practices, complete with a multitude of chapters on materials such as stone, metal, paper, parchment, wood, papyrus, leather, wax, as well as practices such as drawing, chiseling, carving, weaving, etc.

### 7.1.3 Perception and cognition

Physiologically speaking, a visual stimulus must first be seen. More technically speaking, it needs to impinge on the retina of the human eye and must then be transferred to the primary visual cortex at the posterior pole of the brain's occipital lobe. What happens during this process has been described with the help of ERPs (*event-related potentials*). The following description is taken from Rey et al. (2009): In the first 100 to 120 ms after a stimulus has been presented, cognitive activities at a lower level are at work. At this stage, when people see a shape, they realize that something has entered their visual field. Also, they analyze basal features of the visual stimulus, among them contrast, orientation, and connections of lines

246 Cf. https://www.degruyter.com/view/serial/MTK-B?contents=toc-59654 (June 26th, 2020).

(cf. Brem & Maurer 2016: 124). It is also at this point that the presented stimulus is recognized as a special type of stimulus, e.g. a basic shape of Roman script. Further visual processing takes place in neighboring brain areas, predominantly in the ventral part of the occipital and temporal lobes, specifically in the so-called *Visual Word Form Area* (short: VWFA, cf. Dehaene 2009) that is part of the fusiform gyrus (in Brodmann area 37). Between 120 to 180 ms after a stimulus' initial presentation, higher cognitive processes become involved. This cognitive activity is interpreted as the process of recognizing features of basic shapes (in the psychological and cognitive literature, they are called 'letter' features). Crucially, at this point in processing, basic shapes of scripts that exhibit a case distinction are still specified for case, i.e. as lowercase or uppercase basic shapes. It is only after 220 ms that an abstract and case-invariant representation – likely the cognitive representation of what I descriptively termed grapheme (cf. Section 2.2) – is activated. Starting at 300 ms after 'seeing' a grapheme, a person is capable of reacting to the stimulus: they have now perceived the grapheme consciously and can thus follow instructions that require successful prior recognition.

This rough sketch of the grapheme recognition process allows separating temporal stages at which lower cognitive activities are at work from those at which higher cognitive activities are central. The former are led bottom-up by the visual stimulus, while the latter are controlled top-down by graphematic knowledge. (Notably, some models (see below) assume these two types of processes occur simultaneously and not in sequence.) Against the background of this distinction, several interesting questions can be studied, such as: at which point does a stimulus' concrete visual appearance – including specific styles of handwriting or typefaces – play a role? For example, in their study, Keage et al. (2014) showed that in the first 300 ms visual characteristics are crucial. This was tested in an experiment in which participants were presented with different typefaces which had been, based on their visual appearance, categorized as *fluent*, i.e. easy to read, such as Times New Roman or Arial, or *disfluent*, i.e. harder to read, such as Lucida Blackletter. The results suggested that "the initial abstraction of letter meaning is more difficult when the letter is presented in a disfluent typeface, and further, that such a presentation captures more attention than material written in fluent typeface" (Keage et al. 2014: 87). In this section, the focus will not be on how features of concrete stimuli – i.e. concrete materialized graphs – affect processing, but on features of abstract basic shapes that are of relevance. Based on the descriptive structural model of writing established in Part II of this book, a central question that was already partially addressed is: in the course of visual processing, when is a stimulus a *graph*, when does it become a *basic shape*, and when is it finally categorized as a *grapheme*? This question, of course, subsumes the question of when perception becomes understanding, physiology becomes cognition, and graphetics becomes graphematics.

In her (not exactly recent) study, Friedman (1980) broaches this question (cf. also Rothlein & Rapp 2017). In her experiment, participants were presented with – among other things – uppercase and lowercase 'letters' of Roman script (I will stick to her term here; what is meant is 'basic shape'). After the letters had been presented, participants were asked whether they had seen an uppercase or lowercase letter. Strikingly, even in cases in which prior to being presented a letter, participants had been specifically instructed to remember and later state the case in which a letter had been presented, most answers were incorrect: participants could simply not recall the case in which a letter had been presented. Friedman and authors of similar studies assume the existence of an "abstract letter representation", i.e. a combined abstract representation of |A| and |a|, for instance. Transferred to the framework presented here, I argue that these findings either suggest that a grapheme with different allographs has been recognized or that there exists, independently of their assignment to the same grapheme (for whose recognition linguistic information must have been processed), an abstract but not yet linguistic category that subsumes shapes such as |a| and |A| although they differ visually. At a very general level, it is assumed that in the course of the recognition process, graphs are first abstracted to basic shapes and these are then abstracted to graphemes. Graphemes, in psycholinguistic terms, have been called *abstract letter identities* (ALIs, cf. Coltheart 1981: 247; Günther 1988: 156) or *abstract letter units* (ALUs, cf. Finkbeiner & Coltheart 2009: 4). Some authors additionally posit *case-specific letter units* (CSUs, Finkbeiner & Coltheart 2009: 5) which are located at an intermediate level between the concrete visual stimulus (= graphs) and ALUs (= graphemes) and correspond with the descriptive category of basic shapes. Note that from a conceptual point of view, however, the term is a bit too narrow as the same kind of phenomenon can be observed for other pairs of basic shapes such as |a| and |ɑ| or |g| and |ɡ|, which are not distinguished by case (cf. Meletis 2015: 162). If we expand the horizon to other scripts than Roman, we find other distinctions such as the one between positional allographs in Arabic script. These, too, are abstracted at some point in the recognition process and consequently treated as one abstract unit (cf. Carreiras et al. 2013). In a nutshell, it is important to note that the descriptive assumption of the categories *graph* and *basic shape* (at the graphetic level) and *grapheme* (at the graphematic level) are supported by external evidence from processing.

To this day, the *Interactive Activation Model* (IA model) established by McClelland & Rumelhart (1981) remains one of the most influential models of the perception of writing. It incorporates three levels: (1) letter features, (2) letters, and (3) words. The assumption of these levels and their interaction with each other can explain many of the effects observed in recognition and reading processes. This includes, among others, the *word superiority effect* first observed by Cattell (1885) (cf. also Reicher 1969) that describes why people identify individual letters more easily when they are presented as parts of words. The IA model captures that the recognition of units not only works bottom-up, through the identification of features that activate the letters they are a part of, but also top-down, as the recognition of several (not necessarily all) letters already allows an identification of the entire word and this, in turn, makes possible the top-down recognition of the remaining, yet unidentified letters. This bilateral functioning of the model is its greatest strength. One of its weaknesses is its inability to explain letter transposition effects, i.e. when words are recognized although their letters are 'jumbled'. This drawback has prompted the development of newer models that can deal with the positional flexibility of basic shapes and graphemes (cf. Section 7.2.7).

Other influential models striving to explain the reading process include dual-route models such as the *Dual-Route-Cascaded Model* by Coltheart et al. (2001). They rest on the assumption that there exist two possible routes for word recognition: a visual route leading directly to the lexicon and a phonological route that takes a 'detour' through the phonological representation of a word and only then leads to the lexicon. One of the shortcomings of these models is that they are suited to explain recognition processes in English but not in structurally different languages (cf. Christmann 2016: 27; for a general criticism of "alphabetism" in reading science, cf. Share 2008, 2014). Another important model is the connectionist triangle model (cf. Seidenberg & McClelland 1989). Together with similar models, it is subsumed under the group label *Parallel Distributed Models* (PDP models). These models are considered triangular as they assume that orthographic, phonological, and semantic information is processed, and they are parallel in that they claim these different types of information are processed in parallel. In the following, models of recognition will not take center stage but will be referred to at certain points and should be kept in mind, especially with respect to how well they account for graphetic features relevant in processing.

One of the most important but at first glance vague determinants of perceptual graphetic naturalness is visual complexity. Here, complexity refers to the complexity of basic shapes that is evaluated descriptively and that is independent of how it is processed by humans. In an elaborate study, Chang, Chen & Perfetti (2018) offer a useful distinction of factors that contribute to the overall visual complexity of individual basic shapes and, in turn, entire scripts. This multidimensional measure for visual complexity that they call *GraphCom* considers four aspects: first, so-called (1) *perimetric complexity*, which I have called *density* or *grey-scale value* in the proposal of descriptive categories for a script typology. It captures "the density of the written marks ('black ink') relative to the background space in which they are located" (Chang, Chen & Perfetti 2018: 429). In the descriptive graphetic framework presented in Chapter 1, the space relevant for the evaluation of perimetric complexity is the segmental space. As a submeasure of complexity, perimetric complexity is quantitative and size-invariant. This means it is unaffected by the size of basic shapes (or, at the concrete level of graphs, type size or the size of handwritten graphs). As Pelli et al. (2006) showed, perimetric complexity is inversely proportional to the efficiency with which basic shapes are identified. Thus, the greater a shape's perimetric complexity, the greater also the cost for processing. However, perimetric complexity alone is certainly not a sufficient measure of general visual complexity given that it fails to take into account the internal structure of basic shapes. In theory, two basic shapes could share the same perimetric complexity but differ drastically in their structural makeup (cf. Chang, Chen & Perfetti 2018: 430; see below). This is where the remaining three factors of GraphCom come into play. They are highly influenced by *gestalt theory*, especially the principles of proximity, symmetry, convexity, closure, connectedness, and continuation. Accordingly, the second factor contributing to GraphCom is the (2) *number of disconnected components*. For instance, the basic shape |Ξ| of Greek script consists of three unconnected components. Unconnectedness was mentioned above as being unnatural in handwriting because if a basic shape's components are disconnected, users need to lift the writing instrument. Furthermore, visually, the lack of connection between components introduces an empty space at the subsegmental level that might complicate the separation of parts of basic shapes vs. independent basic shapes. The third factor of visual complexity, then, is the (3) *number of connections*. The shape |A|, for example, exhibits three connections between components. These connections have been found to be relevant in the recognition of basic shapes (cf. Lanthier et al. 2009 and below). The fourth and final factor relevant for GraphCom is the (4) *number of components* or elementary forms.[247]

In the study, 133 scripts were analyzed (and *not* 133 "written languages", cf. Chang, Chen & Perfetti 2018: 431). Notably, the authors only considered the form of isolated basic shapes and not the form of sequences of connected basic shapes (cf. Chang, Chen & Perfetti 2018: 432). Moreover, the visual complexity values for different scripts are correlated with different types of writing systems (e.g. alphabets, abugidas, etc.) that the scripts are prototypically employed for. However, the nature of this correlation is explained at no point; instead, the basis of this groundbreaking assumption of a fundamental form-function correlation remains implicit. This is a severe flaw of the study.

Despite its drawbacks, several of the study's findings are striking. Firstly, the number of unconnected elementary forms appears to be most relevant for distinguishing between scripts. Secondly, an additional behavioral experiment confirmed, on the one hand, the hypothesis that two basic shapes that are similar according to GraphCom will also be judged as similar by observers and, on the other, that each of the four dimensions of complexity considered in GraphCom affects processing. In a different experiment, complexity was also correlated with training times needed in a computer simulation of graph learning (cf. Chang, Plaut & Perfetti 2016). Thirdly, "graphic complexity is largely driven by the number of graphs [= basic shapes, D.M.] that is needed in a written language [= script, D.M.]" (Chang, Chen & Perfetti 2018: 438). This means that the number and the com-

**247** Note that Chang, Chen & Perfetti (2018: 449) call these components "simple features (SF)": "SF is a discrete element that can be discriminated from others [...]. Other examples for one simple feature include a line, a dot, a circle, or a curved line". Following Watt (1975), the term "feature" is not fitting for these components; they are rather elementary forms and are themselves made up of abstract features such as [±long] or [±round]. Meshing these two levels together adds redundancy to the level of components as it would make necessary the assumption of various components in cases in which two components share the same shape but differ merely in one feature at a lower level, i.e. orientation, such as |c| and |ɔ| in the shapes |d| and |b| (cf. Section 5.1).

plexity of basic shapes interact in crucial ways. This insinuates that basic shapes of large scripts such as Chinese will exhibit, in sum, a higher visual complexity than the basic shapes of much smaller scripts like Roman (cf. also Section 5.1). Other studies have corroborated the influence of a script's visual complexity on processing and showed that visual complexity affects reading times and knowledge of basic shapes among skilled and beginning readers (cf. Shimron & Navon 1981 for Hebrew, Eviatar & Ibrahim 2004 for Arabic, and Nag et al. 2014 for Kannada).

The visual complexity discussed so far is of quantitative nature. Thus, it evaluates, on the one hand, the quantity of information in general as well as, on the other, different subtypes of information such as elementary forms and connections. There is also a qualitative, structural dimension to visual complexity (cf. Chipman 1977; Chipman & Mendelson 1979). As mentioned above, even when two basic shapes share the same perimetric complexity, number of elementary forms, and (dis)connections, they can still be visually distinct since quantitative complexity reveals next to nothing about the structural makeup of visual information within the segmental space. Quantitative and qualitative visual complexity interact in crucial ways. Most importantly, qualitative complexity can reduce users' perception of quantitative complexity: it is "a structural variable representing organization, symmetry, and other similarity transformations present in the patterns, which reduces perceived complexity" (Gartus & Leder 2017: 19).

In the following, I will first discuss how the force of homogenization (as described by Watt, see above) affects features relevant in perception before shifting the focus onto specific graphetic features that have been found to affect perceptual processing. Together, they contribute to the qualitative complexity of basic shapes. Many of them are well-known as so-called *preattentive features* that are relevant in vision (cf. Wolfe 2000: 344–354) but also as principles of gestalt theory. They include relational features, i.e. features relevant when two or more basic shapes are viewed in relation to one another – such as *similarity*, *distinctiveness*, *redundancy*, and *variation* – as well as individual features that are relevant when analyzing individual basic shapes – such as *symmetry*, *cardinality*, *directionality*, *(dis)-connectedness*, *types of connections*, *location of connections*, and *curvedness*. This list and the ensuing discussion are by no means exhaustive but outline phenomena central for an evaluation of the graphetic processing fit.

At the outset of discussing the graphetic processing fit, Watt's (1983b) four forces of script change were introduced. The force that was claimed to affect perception the most was *homogenization*. The homogenization of basic shapes, i.e. the process in which they become more similar in a script (either graphomotorically or visually) is made possible by perception: shapes that are homogenized first need to be visually perceived. It is also fundamentally driven by cognition, more specifically the human brain's tendency to form systems out of groups of units by rendering these units more systematic. This preference for systematicity leads to a gradual featural convergence of an inventory's units. Notably, although it is ultimately enabled by perception, homogenization is not entirely beneficial for perception. Quite to the contrary, it has a flip side: the introduction of visual ambiguity. In the diachronic development of scripts, homogenization, together with production-driven facilitation, at times resulted in situations in which basic shapes became virtually identical with respect to their visual makeup. An example in which such a *homograph clash* (cf. Brekle 1998: 8) was avoided is the development of the minuscule variant of |G|. Specifically, because of the imminent visual clash with existing |q|, the lowercase equivalent of |G| evolved a left-leaning curved coda in the lower subspace of the segmental space (i.e. in its descender) and became |g|. Notably, this (additional) coda is not located in the central subspace of the segmental space, where codas of lowercase basic shapes are usually positioned, making |g| an exception in the inventory of Roman lowercase basic shapes. Ultimately, this exceptionality results from the need to avoid visual ambiguity with |q| (cf. Brekle 1998: 8). Other prominent examples of changes that led to the collapse of once visually distinct shapes are the basic shapes |ب|, |ت|, and |ث| in Arabic. Their visual convergence was motivated by productional facilitation, highlighting the conflict between productional economy and visual ambiguity. As Salomon (2012: 122) puts it, "excessive simplification carries with it the danger of reducing the visual distinctiveness of a character to the point that it cannot be distinguished readily [...] from other similarly reduced characters". The homogenization of these Arabic basic shapes invoked the counter-force of heterogenization which then led to the secondary introduction of dots so that the basic shapes could be distinguished again (cf. Gruendler 2012: 97f.). Thus, homogenization and heterogenization often go hand in hand in what amounts to a "cyclical process" (Salomon 2012: 123).

As the examples show, homogenization and heterogenization are responsible for two opposing relational principles that affect processing in fundamental ways: *similarity* and *distinctiveness*. The perceptual apparatus tends to favor a certain degree of similarity that is a necessary byproduct of high systematicity in a script. However, before homogenization "overdoes it" (or after it is too late), heterogenization must step in to stop the development or even undo results of prior visual assimilation. Descriptively, the question of how much similarity a script's basic shapes exhibit can be answered by means of the relation of the number of basic shapes and the number of features. Tendentially, if the number of features is greater than the number of basic shapes, sufficient distinctiveness is possible (but must not actually be realized, as not all features might be evenly spread throughout the script), whereas if the number of features equals or is lower than the number of basic shapes, basic shapes might be on the verge of visually collapsing or have already merged. However, this ratio alone is not a reliable indicator of similarity. A second crucial factor is the qualitative nature of the features themselves. In the Arabic basic shapes listed above, for instance, the distinguishing features (the dots) are not as salient as the basic shapes they are added to. Thus, the visual difference between two basic shapes cannot be assessed solely by counting the features in which two given shapes differ but must also include a fine-grained consideration of the nature of these features. |O| and |Q|, for example, differ in one feature, as do |P| and |R|. Yet, one could argue the former are more similar to each other than the latter. Although scripts have yet to be elaborately described in terms of a featural analysis as outlined in Section 5.1, it often appears “visible” to us at first “glance” (in the sense of a naïve perceptual graphetic analysis) that in some scripts, basic shapes exhibit more similarity among each other than do basic shapes of other scripts. As an example, Figure 42 shows a number of easily confusable pairs of Thai basic shapes (cf. Winskel 2010 for an analysis of visually similar basic shapes). Other scripts that show a high degree of similarity among their shapes are Hebrew, the lowercase inventory of Cyrillic, and Armenian (cf. Daniels & Share 2018: 106).

| | | | | | | |
|---|---|---|---|---|---|---|
| ฉ→อ | ฃ→ข | ฑ→ท | ฅ→ค | ต→ด | ผ→ม,น | ย→บ |
| ช→ข | ศ→ค | ส→ล | ฮ→ส,อ | ถ→ด,ก | ซ→ช | ซ→ฃ |
| ฎ↔ฏ | ฬ→พ,ห | อ→ล | ร→ว | ท→ห | ธ→ร,อ | บ→ม |
| จ→ฯ,อ,ฉ | แ→น | ๆ→า | ฯ→ๆ | ไ→ใ←โ | ิ↔ึ↔ื | ั↔็ |

Figure 42. Easily confusable pairs or groups of basic shapes in Thai script, from Punsongserm, Sunaga & Ihara (2017: 17)

Notably, before visual homograph clashes such as the ones mentioned above occur, the force of homogenization actually leads to an increase in systematicity and, thus, to an improvement of a script’s systematic fit. A central process involved in homogenization is overgeneralization. It can be interpreted as a perceptual natural graphetic process because, during processing, it can lead to instances of misremembering. A well-known example of this is |J|, which was mentioned above as an outlier among the uppercase basic shapes of Roman script. If, during production, one feature – most likely the sole feature that does not conform to the feature values of the other basic shapes, i.e. [–right-oriented] – is forgotten, writers (in this case predominantly children in the process of acquiring the script) replace it with the statistically dominant feature or the one that is more natural in the script (for a difference between the two, cf. Section 7.1.1). As a result, children often produce an inverted version of |J| which is not an existing basic shape of Roman script (but, in fact, one of the systematic gaps, cf. Section 5.1). As this example underlines, it is not only productional natural graphetic processes that lead to homogenization but also perceptual ones, even if these also manifest themselves in production.

A completely natural systematic fit, which equals a complete lack of redundancy, can be a disadvantage for the perceptual processing of scripts. An example of an unnatural feature in this context is extrinsic symmetry, i.e. symmetry established between two individual basic shapes in a script. It is the very reason why children have difficulty distinguishing |b| and |d|, for example. To a lesser degree, they are also challenged by horizontal symmetry as in |M| and |W| (cf. Treiman, Kessler & Pollo 2006: 224) or |u| and |n|.

In the case of extrinsic symmetry, several aspects converge to make processing harder. Even though children realize early on that writing is a two-dimensional artifact (see above), one of the central cognitive skills they employ when recognizing three-dimensional objects cannot easily be deactivated in the perception of writing, so-called *object constancy*. This cognitive skill "makes it possible to perceive an object regardless of its orientation in space" (Wiebelt 2004b: 276), which, in turn, makes it difficult for children to accept |b| and |d| as distinctive units because from this perspective, they have the same shape and differ merely in orientation (for a discussion of orientation, cf. also Willows & Geva 1995: 363–365). Extrinsic symmetry, thus, contributes to the similarity of a script's shapes. With the exception of several typefaces in which differences between |b| and |d| are – even if only minimally – emphasized to increase distinctiveness and legibility (cf. Wiebelt 2003: 303f.), they are, in print, always inverted versions of one another. Children, thus, must grasp that the same shape can have different meanings even though they recognize objects (for instance a chair) whether they see them from above, below, or any horizontal angle. Thus, extrinsic symmetry, unlike intrinsic symmetry (which poses no such problems) constitutes a cognitive hurdle that "must […] be overcome during reading acquisition" (Pegado et al. 2011: 742). Further evidence supports the assumption that extrinsic symmetry is unnatural for processing: for instance, it poses problems not only for children in writing and reading acquisition but also for people suffering from disorders of reading and writing (cf. Willows & Terepocki 1993).

Additional evidence comes in the form of the diachronic decrease of extrinsic symmetry in several scripts (cf. Wiebelt 2003, 2004a, 2004b). At the relative beginnings of the history of writing, only intrinsic symmetry was prevalent in scripts. It was the result of the pictographic nature of ancient writing systems, as scribes pictographically manifested the symmetry they perceived in nature and their surroundings in the shapes they produced. In cases in which extrinsic symmetry did occur in these early stages, it was utilized for antonymy, and shapes facing in opposite directions had opposing graphematic values (such as 'come' vs. 'return'). A demonstrative example of the introduction of extrinsic symmetry to a script came with the minusculization of uppercase basic shapes in Roman script (see above). As implied above, complete symmetry was avoided by the introduction of secondary visual features such as serifs. Notably, while the original motivation for serifs might have been productional or aesthetic, their preservation over such a long period of time is arguably due to the benefits they provide for perception (cf. Wiebelt 2003: 303f.). Even after such features that make extrinsically symmetrical shapes more distinctive had been introduced, symmetric shapes represented a challenge for perception. This justifies the question of why extrinsic symmetry was not altogether eliminated. One of the reasons, Wiebelt (2003: 306) argues, is conservatism. Indeed, eliminating or switching out basic shapes that have been integral parts of a script over a long period of time is undeniably an invasive procedure.

In her study, Wiebelt investigated not only extrinsic symmetry in "scripts which have been used for a long time by a large community" (Wiebelt 2004b: 277), which she terms *mature scripts*, but also "scripts which serve special purposes" (Wiebelt 2004b: 277) such as scripts invented for a specific (fictional) work of literature, secret scripts, or scripts intended for a special function. She refers to these latter scripts as *invented scripts*. When comparing the two types, it is striking that invented scripts exhibit a significantly higher degree of extrinsic symmetry than mature scripts. There are various explanations for this. Firstly, extrinsic symmetry, similar to pictography (cf. Section 7.2.2), benefits not only writers who acquire and use an existing script but also the initial creators/inventors of a script: arguably, the practice of creating new shapes by inverting existing shapes requires less effort than having to come up with entirely new shapes. Accordingly, inverting shapes appears to be a productive way of generating new shapes in the context of script creation. Secondly, the factors of continuous usage and the amount of time that has passed play crucial roles: Even if extrinsic symmetry develops at some point due to productional reasons, it is drastically reduced and sometimes even eliminated in mature scripts as they have been used extensively over a long period of time.[248] Invented scripts, on the other hand, are not affected by use and time in the same way. Wiebelt (2003: 321) estimates that with respect to (the lack of) extrinsic symmetry, scripts are perceptually natural if their development has spanned at least 500 years. Finally, as a practical consequence of her study's results, Wiebelt (2004b: 300) cautions that "the creation of a new widespread script should avoid extrinsically symmetrical signs at all costs". This is crucial in the rare instances in which the creation of an entirely new script (vs. the adoption of an existing one) is wished for or necessary.

In sum, many of the types of external evidence central in Naturalness Theory (cf. Section 4.2.4) converge in the case of the unnaturalness of extrinsic symmetry, making it a perfect proposal for a category or parameter constraining the structural makeup of scripts.

Intrinsic symmetry is an entirely different story (for a review, cf. Giannouli 2013). Unlike extrinsic asymmetry, it is an individual feature, i.e. concerns individual basic shapes and not their relation with other basic shapes as part of the same inventory. Thus, it does not pose a cognitive challenge. Concerning anisotropic subtypes of intrinsic symmetry, Morin (2018: 665) claims that vertical symmetry (as in |M|) is more natural than horizontal symmetry (as in |B|) because "[o]ur brains are attuned to vertical (as opposed to horizontal) symmetry". This also echoes the fact that vertical symmetry is the most common type of symmetry found in nature (cf. Wiebelt 2003: 299) and that users become sensitive to vertical symmetry earlier than to horizontal symmetry (cf. Chipman & Mendelson 1979: 375). Notably, intrinsic symmetry has been claimed to reduce the perception of quantitative visual complexity (cf. Gartus & Leder 2017), rendering it a natural feature of writing that stands in stark contrast to extrinsic symmetry.

Extrinsic symmetry was introduced above in the context of redundancy (or the lack thereof) in scripts. To elaborate on redundancy, consider the

248 They are, in Keller's (2014) terms, *phenomena of the third kind* affected by invisible-hand change.

extreme example of Cree script (cf. Figure 28 for a selection of Cree basic shapes). Here, several of the principles sketched above are executed to the fullest possible extent: for example, Cree exhibits the most natural systematic fit with regard to the orientation of basic shapes, as there are no outliers similar to |J| in Roman script. In other words, orientational features – leftwards, rightwards, upwards, downwards – are almost perfectly spread throughout the script. As a consequence, the same shape is consistently used in different orientations, and not just two orientations as in |b| vs. |d|, but all four orientations as in |<|, |∨|, |∧|, and |>|. Crucially, each shape, depending on its orientation, is employed for a different grapheme that in reading needs to be distinguished from the other graphemes exhibiting a different orientation. Specifically, while the orientation-invariant shape consistently corresponds with a consonant phoneme (e.g. all the four shapes in the examples above relate to /p/), the shape's orientation indicates the vowel phoneme. With respect to vowels, the same orientation always indicates the same vowel independently of shape: the right-oriented shapes |<|, |⊂|, and |ᒐ|, for example, all indicate the vowel /a/. Thus, both shape and orientation are diagrammatical features. From a semiotic point of view, this is highly natural (cf. Section 6.2). However, because of the above-mentioned cognitive skill of object constancy, the status of orientation as a graphematically distinctive feature poses a challenge for processing. As Treiman & Kessler (2014: 166) note, if a child acquiring literacy in Cree were to misremember a certain grapheme's orientation when writing or reading, filling the feature in with the wrong value, i.e. the wrong orientation, will in all cases result in the production or perception of a wrong shape, which automatically leads to the production of a wrong grapheme. Ultimately, this results from the lack of redundancy with respect to the feature of orientation, as there are no systematic gaps, i.e. every basic shape that is imaginable is actually an existing unit of the script. If there were, by contrast, a certain degree of redundancy, misremembering a (feature of a) given basic shape would also result in the production of an 'erroneous' basic shape. This erroneous shape, however, might not be an actual unit of the script but instead a systematic gap that is not associated with any grapheme (such as |ᒐ| in Roman script).

As a relational phenomenon, redundancy is not only affected by potentially distinctive features (such as orientation) but also by features that are non-distinctive in a script. In this vein, visual variation at the concrete level of graphs, i.e. the concrete materialization of scripts, also represents a form of redundancy. From a sociosemiotic and sociocultural perspective, expressiveness is a resource of writing because it allows people to express facets of their personality or the specifics of the writing situation – in a nutshell, to position themselves socially with the help of graphetic resources and variation. At the same time, however, from the perspective of processing, graphetic variation presents a challenge to beginning readers who must learn to separate distinctive, i.e. graphematically relevant features, from non-distinctive (e.g. decorative) features (cf. Treiman & Kessler 2014: 161f.). An early stage in this process is reflected by the initial over-emphasis of distinctive features in literacy acquisition, e.g. when children produce the dots in Arabic basic shapes much bigger in relation to the rest of the basic shape than is common (cf. Kenner 2004: 83f.; cf. Figure 43).

Unlike the problems posed by the similarity of distinctive basic shapes in a script, the cognitive hurdles of visual variation, however, are soon overcome. Interesting evidence on visual variation comes from the relationship between the production and perception of handwriting: studies suggest that the acquisition of scripts with higher visual complexity strengthens visual skills (cf. Chang, Chen & Perfetti 2018: 429; Kenner 2004: 76). In particular, the fact that children produce variable graphs in handwriting, rendering their written output visually variable, appears to benefit the perceptual system because it receives a multitude of perceptually variable exemplars of one category (or tokens of one type), i.e. multiple graphs for one basic shape (cf. Li & James 2016). Notably, the graphs that children produce are themselves also dependent on abstract mental representations of basic shapes that have been established based on prior perception.

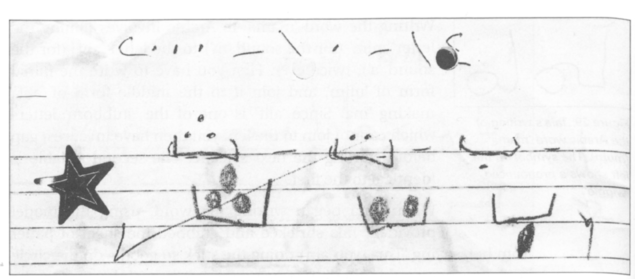


FIGURE 43. Overemphasis on dots in Arabic, from Kenner (2004: 84)

Redundancy can not only be interpreted as a relational feature relevant for the analysis of scripts but also as an individual feature important for the analysis of single basic shapes. In this vein, the results of Changizi & Shimojo's (2005) study suggest that in many of the world's scripts, already half of the features of basic shapes would suffice to recognize them. However, features or, more generally put, visual informational load, is not distributed evenly in basic shapes. This was shown by Kolers (1983: 373) who observes that "not all parts of a letter contribute equally to its identity". He speaks of a *polarization* of information: specifically, cutting away either the right halves or the left halves of Roman uppercase and lowercase basic shapes reveals that information more relevant for the recognition of shapes appears to be 'polarized' in their right halves. This, Kolers (1983: 376) hypothesizes, might be motivated by eye movements, particularly their direction: the reading direction in writing systems using Roman script is dextrograde, meaning the eyes move rightwards, and this is where the bulk of the relevant visual information is positioned. By contrast, for Hebrew shapes, employed in a sinistrograde writing system, Kolers' informal analysis reveals that information is polarized in the left half, and in Chinese characters, the bottom half as well as the left side appear to be more informative (cf. Chou 1930), coinciding with the traditional sinistrograde downward reading direction. This directional polarization, of course, could simply reflect the parameter of writing-direction orientation that was outlined above and favors codas and appendices in basic shapes that are located in the writing direction.

Another question relevant for graphetic processing naturalness is whether elementary forms within basic shapes are cardinal (such as |I|) or oblique (such as |/|). Cardinality was the focus of a study by Morin (2018) who hypothesizes that "cardinal (vertical and horizontal) orientations, being easier to process, should be overrepresented in letters" (Morin 2018: 664). Here, easier to process means easier to recognize, discriminate, and memorize (cf. Morin 2018: 665). Morin's analysis of 116 scripts shows that cardinal elementary forms are not only overrepresented in visual systems but also tend not to mix with oblique strokes. Thus, mixed basic shapes such as |K| are rare. In fact, Roman script exhibits a separation effect in that its basic shapes are either composed purely of cardinal segments such as |E| or oblique segments such as |W|. Mixed shapes such as |K| as well as shapes consisting only of oblique elementary forms (such as |W|) are less natural specifically because oblique segments often cause acute angles which are, for example, less faithfully copied by children than right angles (such as in |T|, cf. Davis, De Bruyn & Boyles 2005). Fiset et al. (2009) confirm that in addition to line terminations (see below), vertical strokes appear to be the most important feature in the recognition of uppercase Roman basic shapes.

As established above, the number of elementary forms as well as their connections and disconnections in basic shapes are relevant factors of quantitative visual complexity, whereas the exact nature of elementary forms is a matter of qualitative visual complexity. Concerning the latter, Changizi et al.'s (2006) study shows that several visual topological configurations prevalent in nature are reflected in the structural makeup of basic shapes. A highly relevant variable in visual processing is topology (cf. Wolfe 2000: 353f.), understood as the spatial arrangement of elementary forms within basic shapes. This subsumes not only the connections between elementary forms but also their individual positions within the segmental space as well as their spatial arrangement with respect to one another. There is ample evidence that in perception, global topological configurations are used to distinguish between shapes (cf. Kanbe 2013). For example, it was shown that the removal of internal parts of strokes is less detrimental to processing than the removal of connections, i.e. apices and vertices (cf. Lanthier et al. 2009 and Figure 44), highlighting the relevance of topological information in the form of connections. The visual complexity of different types of connections has been quantified (in theory): one proposal assumes continuous contacts (as in |O| or |~|) to be least complex, followed by crisp contacts as in |T| or |F| and crossings (such as in |X| or |+|), which are claimed to be most complex (cf. Altmann 2004: 69f.). However, research on the different effects these connections have on processing is scarce. As mentioned above, line terminations have been found to play a crucial role in visual processing: the lower line termination in |C|, for example, makes it possible to distinguish it from |G| and |O| (cf. Fiset et al. 2008: 1166f.). |C| and |O|

differ also in closedness (corresponding with the gestalt theoretical principle of closedness), which, in addition to connectedness, has been described as an important feature for visual processing (cf. Chen 1982: 699).

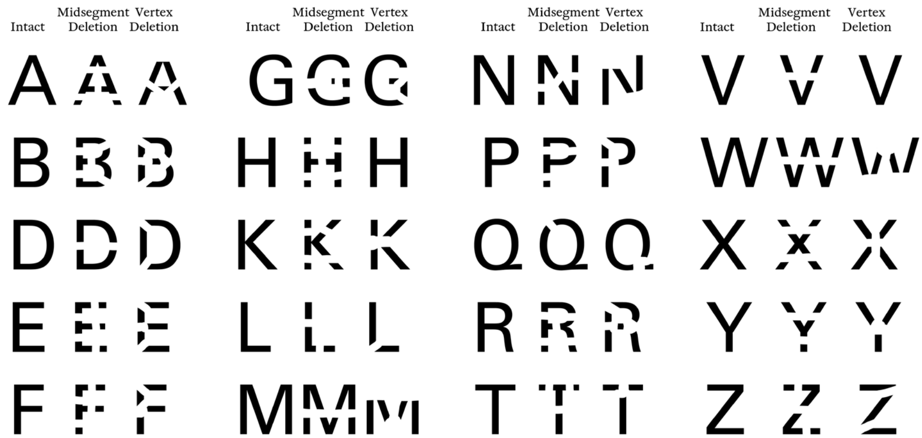



FIGURE 44. Deletion of midsegments and terminations, from Lanthier et al. (2009: 68)

To study connectedness as a relational rather than an individual feature, it is revealing to consult research on connected scripts such as Arabic as well as to compare unconnected print with cursive connected handwriting in scripts such as Roman. While connectedness was identified as a natural feature in production, is the same true for perception, or is the lack of visual segmentality detrimental to the perception of individual shapes? For writing systems using Roman script, whose shapes appear unconnected in the vast majority of typefaces, evidence points to the latter. In a study involving French adults and children, both groups recognized cursive words with connected graphs less efficiently than printed words with unconnected graphs (cf. Danna et al. 2018). This was also confirmed for Spanish, where "a small, but significant advantage in response times for those words composed of separated letters than for those composed of connected letters" was observed (Roldán, Marcet & Perea 2018: 285). Interestingly, in Arabic, the opposite appears to be the case: several studies found that connected words are more easily processed than words in which graphs are unconnected. In other words, connectedness benefits the reading process (cf. Taha, Ibrahim & Khateb 2013; Khateb et al. 2013). These diverging results are relevant concerning the interaction between the subtheories of naturalness: while it is certainly possible that either connectedness or segmentality are universally natural graphetic features, users of Roman script are more familiar with unconnected writing, users of Arabic script more familiar with connected writing. These respective features override the universally natural feature, making them more natural for users at the system-specific level (cf. Khateb et al. 2013: 226f.). This does not settle the question of whether segmentality or connectedness are universally more natural. However, the fact that in Arabic, too, there are empty spaces between words, suggests that segmentality is to some degree natural also in Arabic. Word spaces guide eye movements and the reading process even in the absence of spaces between graphemes (cf. Section 7.2.8). A writing system that lacks segmentality entirely would, probably, even if connectedness is natural at the system-specific level, be harder to process than a writing system that does exhibit segmentality at the grapheme or word level.

A final perceptual feature worth mentioning is curvature (or roundness), the opposite of angularity. Not only individual basic shapes within a given script differ in curvedness (such as |H| and |C| in Roman script) but also entire scripts: the Georgian and Burmese scripts are very curved when compared with the angular Hangul or Hebrew scripts, for example. The question of whether round or angular basic shapes are more natural perceptually has (to my knowledge) not yet been studied, although various types of evidence point to a preference for curves. For example, curves are aesthetically preferred (cf. Cotter et al. 2017; Bar & Neta 2006; Silvia & Barona 2009). This was shown by means of typefaces that were adjusted to be more rounded or angular (cf. Kastl & Child 1968). In a similar vein, Velasco et al. (2015: 1) instructed participants to match round or angular typefaces with taste words and found that round typefaces were associated with attributes such as "sweet", while angular typefaces were associated with "bitter", "salty", and "sour". The authors hypothesize that this could reflect the fact that round typefaces are easier to process. Note that the causal relationship might also be inverse: aesthetic preferences and attitudes towards typefaces might affect the ease of processing. Actual evidence on how these features affect processing is almost non-existent; only bits can be discovered scattered throughout the literature. For instance, a study that focused on visual processing in dyslexics used basic shapes from Georgian script, which are predominantly round, and found that dyslexics performed just like normal readers (cf. Shovman & Ahissar 2006). While the visual appearance of the shapes was not specifically tested – instead, factors such as letter size, letter crowding, the addition of white noise, etc. were analyzed – the results of the study indicate at the very least that roundness is not detrimental to reading. In fact, there might not be a preference for curvedness or angularity in perception at all. As several studies indicate, both curves and round shapes as well as straight lines and angular shapes are relevant in the perception and recognition of basic shapes (cf. Chang, Furber & Welbourne 2012: 2786; Fiset et al. 2008, 2009). Undoubtedly, curvedness and angularity are two of the (many) features for which more evidence is needed before they can be reasonably included in a comprehensive functional theory of writing.

## 7.2 Graphematics

From the perspective of Naturalness Theory, on which this sketch of a functional theory of writing is based, and specifically according to Natural Morphology, the graphematic processing fit is arguably at the very core of an investigation of what is natural in writing. Essentially, it identifies the features that prove crucial in how users process writing – and not just its substance (which is studied by the graphetic processing fit, see the previous section) but mainly its linguistic functions. In Natural Morphology, structures that are easy to process for the human brain were defined as natural, with the core assumption being that facets of semiotic structure have a direct bearing on how easy or difficult it is for users to process signs. In the functional theory of writing that is proposed here, these two aspects – structure (determined descriptively) and processing (determined with external evidence) – that were originally interpreted as being causally linked in a direct and inseparable manner are carefully separated into the linguistic fit on the one hand and the processing fit on the other. This separation is necessary as it is one thing to analyze descriptively whether the relationship between a language and the graphematic module of its writing system is, for instance, transparent, and another thing to investigate whether a transparent graphematic module (such as Italian) is processed with less effort than a less transparent graphematic module (such as German). Intuitively, it appears self-evident that the linguistic fit affects the processing fit, i.e. that the way a "language is written may matter for reading [and writing, D.M.] processes" (Perfetti & Harris 2013: 296), but this is not a strong enough argument against a separation of these two aspects (at least in the first step). Furthermore, the influence of one fit on the other is not unidirectional: diachronically, it is claimed that the processing fit (i.e. 'human pressure') has strongly shaped the structure of writing (cf., for example, Dehaene 2009). Consequently, the linguistic fit of writing systems that can be evaluated synchronically results at least in part from the (suboptimal) processing fit of their previous developmental stages that exhibited less natural linguistic fits. Finally, since writing is a cultural technique, the sociocultural fit (cf. Chapter 8) is so powerful that in some contexts, it has the potential to override both the linguistic and processing fits.

First, it must be said that several factors complicate the analysis of the graphematic processing fit. Firstly, the lack of pertinent data. Granted, the claim that there are only little relevant data concerning the cognitive and psychological processing involved in writing and reading is false. In fact, such data are strikingly abundant, especially in the realm of psychology. However, on the one hand, this abundance makes filtering and identifying the literature that is truly relevant difficult – especially for a linguist or generally a scholar who is an outsider to the pertinent fields. On the other hand, the perspective taken in much of the interdisciplinary research challenges its operationalization for the processing fit: the majority of studies interested in reading, writing, and the acquisition of both of these (bundles of) processes focuses on the cognitive skills or abilities they require instead of how they are influenced by properties of writing systems (i.e. the linguistic fit).

A central notion worth mentioning as an example is *phonological awareness*. Although there is no real consensus on a definition, phonological awareness is the subject of countless studies. Interestingly, while its definition frequently remains vague, this does not keep scholars from investigating its effects on reading and reading acquisition (and, although not as often, spelling and spelling acquisition). A great number of psychological comparisons of writing systems thus explores the question of how phonological awareness influences reading acquisition and literacy development in the studied writing systems. In the process, however, the structural properties of these writing systems themselves are mostly relegated to the background. Research on reading and writing thus largely dismisses the fact that as a cognitive skill, phonological awareness is crucially dependent on and interacts with the (semiotic) structure of the writing system that is being acquired. Indeed, structural differences are the very basis for the fact that phonological awareness differs in children who become literate in distinct writing systems. Admittedly, when seen from the inverse perspective, linguistic research on the structure of writing systems often discards psychological and cognitive implications (but cf. Daniels & Share 2018 for a paradigm shift). In short, the big challenge of outlining the processing fit is to combine relevant findings from the many grapholinguistic subdisciplines, which first requires to look at them from different perspectives than the ones they were originally intended for.

The following preliminary treatment of the graphematic processing fit takes the properties of writing systems – more specifically, features of their graphematic modules – as a basis and investigates the way they influence how users process writing. The parameters described in the context of the linguistic fit (cf. Chapter 6) serve as a starting point. For each parameter, several questions will be addressed: (1) Is the parameter relevant in the processing of writing systems? If so, how? (2) Are the values of the parameter that are descriptively natural (i.e. for the linguistic fit) also natural regarding processing needs? In other words: are the parameter's linguistic and processing naturalness values congruent or incongruent? (III) What is the precise interaction of the linguistic and processing fits of the parameter? Does (or has) the processing fit influence(d) the linguistic fit, supporting the view that human pressure shapes the development of writing systems?

Notably, in the discussion of the graphetic systematic and processing fits, the naturalness of the grapheme's signans – the basic shape – was assessed. Here, the other component of the sign, the signatum (i.e. the linguistic unit), will be highlighted. While evaluating how natural the individual linguistic units (such as specific morphemes) that serve as signata are would exceed the scope of the present study (and is actually the subject of other subbranches of Naturalness Theory such as Natural Phonology and Natural Morphology, cf. Section 4.2), what can be reasonably discussed and what has indeed been the topic of past grapholinguistic discussion is *what type* of linguistic unit (such as 'the morpheme') serves, with respect to the processing fit, as the most natural basis of writing systems at the universal level. A candidate for such a most natural *type* of signatum that has been suggested is the phonological syllable. The following section will critically sketch and evaluate the discussion that led to this assumption before the other naturalness parameters will be treated in subsequent sections.

### 7.2.1 Unit of processing

To establish which unit of language has the most natural linguistic fit for a given writing system, both the type of the language the writing system is based on and the idiosyncratic features of said language must be considered (cf. Section 6.1). Notably, while a unit of representation with a good linguistic fit is natural descriptively, it does not necessarily also have to be the unit that is most natural for processing. Indeed, evidence suggests that one type of linguistic unit might be universally preferred in processing, i.e. even across typologically diverse writing systems: the (phonological) syllable.

Evidence that underlines the naturalness of the phonological syllable for processing comes in a striking number of different forms of evidence relevant in Naturalness Theory (cf. Section 4.2.4). For instance, the history of writing provides strong indications: Daniels (1992) illustrates a possible primacy of the syllable by means of ancient grammatogenies, i.e. the initial geneses of writing which occurred for/in Sumerian, Chinese, and Mayan. These grammatogenies were, in his terminology, *unsophisticated*, meaning that the cultures in which these writing systems were created had been oral, i.e. were not literate prior to the creation of the writing systems. Crucially, this means that the members of these cultures had not been familiar with the concept of literacy, either because it did not yet exist (at all) or because they simply were not aware that it had already been invented in other cultures. Put differently, these cultures introduced not only writing systems but writing and literacy *per se*. Strikingly, now, in all of these first instances of writing, the result was a morphosyllabographic writing system. For Chinese, Daniels (1992: 83, 94f.) does mention the possibility of "stimulus diffusion" but then deems it unlikely that the Chinese borrowed the idea of writing from Sumerian. By comparison, modern grammatogenies, i.e. "inventions" or introductions of writing (systems) in a given culture, can either be unsophisticated (like the mentioned ancient grammatogenies) or sophisticated. In the latter case, cultures borrow the idea of literacy from another culture. To provide examples of both, the creations of the writing systems of Cherokee and Vai were unsophisticated, whereas the conceptions of Korean Hangul, which is syllabically spaced (see below) or the Cree syllabary were sophisticated. These examples imply that both in ancient and modern grammatogenies as well as in unsophisticated and sophisticated ones, the syllable takes on a central role. Unsophisticated grammatogenies, in which writing systems were created "from scratch", are of course most striking. Notably, in the ancient grammatogenies mentioned above, the choice of the syllable as the unit of representation could be determined by language-specific properties: in all languages in question, i.e. Sumerian, Chinese, and Mayan, words were typically monosyllabic (cf. Daniels 2017: 84). The convergence of syllables and words actually calls into question whether it was truly the syllable that was central in these grammatogenies or whether the word, arguably also a very salient unit, was more important (especially since it is, from the perspective of the double articulation of language, the level of primary articulation). In this context, it must also be noted that the word is the more easily accessible unit given that it does not require users to possess the same amount of metalinguistic awareness as the secondary level of (syllabic) phonological representation. This possible salience of the word might be one of the reasons that Daniels' "universal" of "(mono)syllabic origins for writing" has been called into question (Klinkenberg & Polis 2018: 59, 93), although neither specific reasons that justify the doubts nor counterarguments have been presented.

Daniels (2017: 76) cites further evidence arguing for the primacy of the syllable from fields such as psycholinguistics, literacy instruction, and phonology:

> Psycholinguists find that people not literate in an alphabetic script are unable to manipulate portions of the speech stream at the level of the segment [...]; educational psychologists find that syllabic approaches to teaching children to read can be more successful than approaches requiring them to identify subsyllabic segments [...]; phonologists increasingly work with levels of analysis other than that of the segment or individual sound [...].

The first type of evidence mentioned here is phonological awareness, more specifically the distinction between syllable vs. phoneme awareness. Phonological awareness is *per se* a problematic concept (see above). There is ample evidence pointing to the special status of the phonological syllable that stems from studies primarily focused on writing (rather than studies that investigate phonological awareness specifically), and in the following, the focus will be on them. This evidence does, however, often allow drawing more general conclusions about phonological awareness.

FIGURE 45. Subset of Cree basic shapes used in the study, from: Inkelas (2013: 79)

A study that concerns the second type of evidence mentioned by Daniels – it could be called 'educational evidence' – is Inkelas et al. (2013). It tests the 'learnability' of writing systems, more specifically the question of whether "the acoustic stability of the speech chunks mapped to symbols is a factor in subjects' ability to learn a novel writing system" (Inkelas et al. 2013: 75). Four types of units of representation were compared with respect to their learnability: (1) segments, i.e. phonemes, (2) moras, i.e. for each CV and for each consonant coda there is a grapheme, (3) onset-rimes, i.e. for each consonant onset and for each VC rime there is a grapheme, and (4) demisyllables, i.e. for each CV and VC there is a grapheme. Participants had to memorize twenty basic shapes that were taken from the larger inventory of Cree basic shapes (cf. Figure 45). In the context of the study, these shapes were graphematically related to one of the four above-mentioned types of units, and roughly an equal number of participants was assigned to each type. In advance, the authors had formulated two conflicting hypotheses: firstly, the *acoustic stability hypothesis* predicting that graphemes that represent acoustically stable speech chunks are learned more accurately, i.e. that learnability decreases in the order demisyllable > mora, onset rime > segment. Secondly, the *alphabetic familiarity hypothesis* predicting that native speakers and readers of English are biased towards an alphabetic, i.e. segmental system, and that this influences the study's results.

Participants were visually presented the basic shapes while their respective graphematic values, i.e. their 'pronunciations', were played to them as prerecorded sound files. After the completion of this learning stage, a phase of combination training followed: here, participants were trained on how to combine the graphemes they had previously learned to form CVC words. Finally, after the learning and combination training phases, the actual testing commenced: subjects were asked to read aloud novel CVC words combined from the graphemes they had acquired and trained with in the previous stages. For example, they were instructed to read the word /gik/, which was graphematically represented in different ways across the four conditions: as three different graphemes <g>, <i>, <k> in the (1) segmental condition, as two graphemes <g> and <ik> in the (2) onset-rime condition, as <gi> and <k> in the (3) mora condition, and as <gi> and <ik> in the (4) demisyllable condition. The results confirm the acoustic stability hypothesis, as participants "learn a writing system better if the symbls [sic] are presented to them in speech chunks larger than the individual phone" (Inkelas et al. 2013: 88). While a single study is evidently not conclusive, this interesting experiment provides further evidence that – at least in the context of literacy acquisition – larger phonological units might be more natural processing units than segments.

Evidence for the importance of the syllable also comes from production. For example, studies have pointed to the importance of syllable boundaries in handwriting. Accordingly, Kandel, Álvarez & Vallée (2006) found that the duration of intergrapheme intervals was longer between syllable boundaries than within syllables. In three experimental conditions, French and Spanish participants were tested. A striking observation was that the interval between the graphemes of the sequence <gn>, which is always intrasyllabic in French (such as in <consi.gner>) and intersyllabic in Spanish (<consig.nar>), was shorter in French than in Spanish. This was true not only for monolingual French or Spanish writers but also for bilingual French-Spanish writers, who systematically produced a shorter interval when writing French (cf. Kandel, Álvarez & Vallée 2006: 26).

A question that is once again raised in this context concerns the correspondence between the graphematic syllable and the phonological syllable. In French, as in German (cf. Section 2.4), the two are not always congruous: for example, <bar.que> is bisyllabic, whereas its phonological representation /baʁk/ is monosyllabic. When the production of words with incongruous phonological and graphematic syllabifications is compared with the production of words in which the two types of syllables correspond (such as <bal.con>, /bal.kɔ̃/), it becomes evident that even in phonologically monosyllabic words (as in /baʁk/), a purely graphematic syllable boundary (as in <bar.que>) has a bearing on processing, as suggested by letter stroke duration and handwriting fluency (cf. Kandel et al. 2009). This lends strong external support for the descriptive assumption of an independent graphematic syllable (cf. Section 2.4). Notably, the graphematic syllable is not only relevant in handwriting but also in typing: Will, Nottbusch & Weingarten (2006) showed that inter-keystroke intervals were longer at syllable boundaries than intrasyllabically, and they were even longer at syllable boundaries that are simultaneously morpheme boundaries. The authors take this as evidence that the motor system works with sublexical units instead of holistic words. Strikingly, this not only applies to hearing but also to deaf writers, whose inter-keystroke intervals were also longer at syllable boundaries (cf. Nottbusch et al. 2005). This delay in deaf writers can obviously not be explained by phonology, which serves as further evidence for the independence of a graphematic syllable in processing.

In general, what must be mentioned here is that a focus on the concept of phonological awareness and, thus, the phoneme and phonological segmentality, is likely (yet) another reflection of alphabetocentrism. When typologically different writing systems are considered, a different picture emerges: recent research on Asian writing systems, for example, has deemphasized the role of the phoneme in processing in these systems while highlighting the relevance of the syllable (cf. Winskel 2014: 174).

A final remark concerns the fact that even in writing systems that are not syllabographic, a level of syllabic representation can take on a structurally relevant role. In the case of the graphematic syllable that was postulated for German (cf. Section 2.4), for example, it is the visual feature [+length] exhibited by certain basic shapes that visualizes graphematic syllables (cf. Fuhrhop & Buchmann 2016). As was argued above, a similar principle can be found also in other alphabets such as Greek, Georgian, and Armenian. In Korean, too, graphemes themselves are segmental, and yet they are arranged in syllable blocks. This renders syllables visually and graphematically more salient than graphemes. Against the background of the hypothesis that human pressure shapes the makeup of writing systems, the recurrence of syllabic structures in writing systems (regardless of their type) could be a result of the syllable's natural processing fit.

### 7.2.2 Iconicity

In the context of evaluating the linguistic fit of the graphematic module, I argued that pictography is to a certain degree unnatural because it is impossible to pictographically represent all of a language's elements, which, however, need to be represented in a full writing system. The elements that are hard or even impossible to represent include, among others, abstract concepts and function words. The naturalness of iconicity that was postulated in Natural Morphology is based on the benefits it is claimed to offer for processing, particularly perception, as signantia that bear a resemblance to the signata they are semiotically linked to are claimed to be easier to recognize (cf. Section 4.2.3). This is certainly true for pictography in writing as well. However, while pictography might be natural for perception, it requires high cognitive and physiological effort in production; indeed, pictographically representing the objects/referents that linguistic units are related to is an endeavor not dissimilar to drawing (cf. Tversky 1995: 34f.). Thus, production and perception are in conflict. Notably, pictography, and iconicity in general, must be understood as a matter of degree, of course. A highly pictographic grapheme is likely harder to produce than a slightly pictographic grapheme whose basic shape resembles the object referred to by the signatum not as straightforwardly.

In short, in terms of processing, the drawbacks of pictography appear to outweigh its benefits. This fits the plain observation that modern writing systems exhibit only small degrees of pictography. Diachronically, however, writing systems (and strikingly, each of the independent inventions of writing – Sumerian, Chinese, and Mayan) were characterized by remarkable pictography. This divergence between ancient and modern writing systems or, from a dynamic point of view, the gradual historical decrease of pictography suggests that the unnatural aspects of pictography dominated over its natural aspects which were the basis for its initial prominence in (the early stages of) writing systems.

With respect to this reduction of pictography, recent studies in the context of cultural evolution prove revelatory. Even though they do not deal with writing specifically, they address aspects that are still relevant to the evolution of writing. Two of them shall be described in more detail. The first of them is a study by Garrod et al. (2007) in which participants were instructed to engage in a task that resembled the game *Pictionary*: one player draws a picture of an object, person, concept, etc. while the other player must recognize what is being/has been drawn. The study consisted of three experiments of this kind.

(1a) In the first experiment, it was shown that the repeated drawing of an object is insufficient for the drawing to become less pictographic and overall simpler. On the contrary, when a person repeatedly drew a concept without receiving feedback from anyone until after the drawings had been completed, "the drawings became increasingly complex [...] and retained their iconic character" (Garrod et al. 2007: 983). This condition was the so-called *SD-F condition* (for 'single director without feedback').

(1b) By contrast, in the *SD+F condition* ('single director with feedback'), the drawer received feedback by a so-called matcher who had not directly observed the drawing process but was shown the drawing(s) after each of the six blocks (i.e. repetitions) and was instructed to give feedback. This feedback, in turn, allowed the drawer to adapt the drawing in each new block according to the feedback that the matcher had given following the preceding block.

(1c) In the third and final condition of the first experiment, the *DD+F condition* ('double director with feedback'), the two participants alternated roles over six blocks: after one participant had been the drawer in the first block, they became the matcher in the second block. Like in the SD+F condition, in the DD+F condition, too, the two participants were separated visually while the drawing was taking place and the matcher was only allowed to give feedback *after* the drawing had been finished. The feedback, thus, was non-concurrent.

(2) In the second experiment, Garrod et al. tested whether the *quality* of feedback has a bearing on the 'evolution' of how much pictography drawings exhibit. This was done by comparing the above-described DD+F condition with a similar condition in which, however, the feedback was concurrent: the two participants were standing side by side whenever one of them was drawing (cf. Figure 46 for an example of drawings produced in this condition). Crucially, thus, the matcher could give live feedback while the drawing process was taking place.

(3) Finally, in the third experiment, 'overseers' were given the drawings produced in the SD+F and DD+F conditions of the first experiment. These overseers had not participated in the first two experiments. Two types were distinguished: *early overseers* were given all of the drawings of a given concept, i.e. all the drawings from blocks 1 to 6, while *late overseers* were given only the drawings that were produced from blocks 4 to 6. The authors hypothesized that, in general, overseers should perform significantly worse than matchers who had participated in the original experiment since overseers were not able to interact with the producers of the drawings. A second assumption was that late overseers should perform even worse than early overseers because they were not given the full history of a drawing's development.

The results of the three experiments are striking. In both the SD+F and DD+F (without concurrent feedback) conditions, drawings became simpler (= composed of fewer elements) and less pictographic over the course of the six blocks. The fact that participants in these conditions could receive and give interactive feedback following each block evidently served as a form of communicational interaction that allowed *symbols* (as opposed to pictographic *icons*) to arise, i.e. signs that bear no visual resemblance to the concepts they represent. Additionally, in the DD+F condition with concurrent feedback in which both participants alternately drew with the other one observing the drawing process, the drawings of the two participants became gradually similar. What does this mean for the naturalness of pictography in the processing of writing?

Block 1 (CF) Block 2 (CF) Block 3 (CF)

Block 4 (CF) Block 5 (CF) Block 6 (CF)

FIGURE 46. Example of the development of the drawing 'computer monitor' in the DD+F condition (with concurrent feedback) from Experiment 2, from Garrod et al. (2007: 978)

To start at the beginning, the initial presence of pictography in a writing system can (likely) be explained by two factors: (i) the creation of basic shapes that visually resemble what they signify arguably requires less effort than devising abstract and arbitrary basic shapes, at least for basic shapes that graphematically relate to concepts referring to concrete objects. Note that the restriction to concrete objects makes pictography (relatively) easier to produce only in morphographic systems (but cf. also Hangul for a phonographic example, Section 6.2). Secondly, (ii) pictography has benefits for perception, i.e. recognition and reading. While the initial creation of a writing system as well as the ongoing perceptual processing on behalf of its readers are facilitated by it, pictography challenges writing processes. As shown in Section 7.1, production is driven largely by the force of facilitation. In this context, pictography and articulatory simplicity are in conflict, with simplicity apparently winning out in the end. This leaves open the question of *how* exactly pictography is diminished. Garrod et al.'s study provides a promising preliminary answer to this question: it is communicational interaction that allows the rise of symbolicity and the associated establishment of symbolic conventions. Accordingly, the decrease of pictography is driven not only by the force of facilitation that acts upon the production processes in individuals: indeed, as the SD-F condition in Garrod et al.'s (2007) Experiment 1 showed, without feedback, repeated drawings of the same concept do not become simpler. Furthermore, as the results of Experiment 3 suggested, participants who had not been involved in the initial experiment had noticeable difficulty recognizing the concepts that the drawings (especially those produced in later blocks) represented. This is because they had not taken part in the drawings' development and therefore lacked knowledge of the graphical history of the drawings. In short, while the decrease of pictography may ultimately be conditioned by the facilitation of production, it is the process of establishing conventions that makes possible the ensuing symbolicity and arbitrariness prevalent in writing systems.

In other words, the production of shapes in a community whose members continuously use them and, crucially, use them in communication with other members, allows for conventions to be established, a process referred to as *conventionalization* (cf. Tversky 1995: 35). Consequently, basic shapes or graphemes[249] that were once pictographic become abstract and arbitrary. Of course, the actual continuous use of a writing system in 'real life' differs drastically from the experimental settings in Garrod et al.'s study. However, based on their results, the authors suggest "there may be an alternative grounding mechanism that could operate with repeated and reciprocal use of graphical signs in a community even when there is no direct interaction (e.g., as in the use of a writing system)" (Garrod et al. 2007: 983). By that, they imply that in the actual use of a writing system, feedback of some kind is offered by other users of the system, too, even when they are not present at the time of production (as was the case in the experimental DD+F condition) or do not give feedback right after production (as in the SD+F condition). When writing is used for communication, which was its primary function in the pictographic stages of writing systems such as Sumerian, Mayan, and Chinese, scribes write to be understood, making it likely that even in the use of these ancient writing systems there existed some sort of feedback to indicate to the producer whether communication had been (un)successful.

Ultimately, pictography is neither an absolutely unnatural nor absolutely natural feature of writing, as this depends on the perspective taken. An interesting idea that Garrod et al. (2007: 963) formulate is that "during the evolution from iconic to symbolic graphical representation, structural complexity migrates from the sign to memory representations in sign users". This emphasizes the tradeoff between physiology and cognition. When shapes become easier to produce (and, at a merely physiological level, to perceive), they require increased cognitive effort to be recognized and categorized since the relations between visual and linguistic units become increasingly arbitrary. However, as Tamariz et al. (2018: 347) note, "[g]raphical iconicity is costly in terms of time and effort – the details need to be drawn accurately when the conventions are being established, but once the conventions are entrenched, simplified forms work just as well". This implies that while there might be more cognitive effort needed to acquire and memorize arbitrary graphemes, once they are mastered, they are not necessarily less suited to processing than pictographic graphemes.

The second study worth mentioning is Caldwell & Smith (2012). While Garrod et al. (2007) investigated *dyads*, i.e. groups consisting of two partic-

**249** Pictographic basic shapes and pictographic graphemes are not the same. A shape itself, independently of what linguistic unit/information it might relate to, can be iconic semasiographically when it directly resembles a thing or concept. A grapheme, on the other hand, is pictographic only when the shape is in an iconic relationship with the signatum, i.e. the linguistic unit that is also a part of the grapheme. These two types of iconicity can coincide such as in the Chinese character for *tree*, where the shape visually resembles a tree and relates to the morpheme with the signatum 'tree' (cf. Section 7.1.1).

ipants each, Caldwell & Smith observed the behavior of so-called *replacement microsocieties* (see below), arguing that studies such as Garrod et al. (2007) "imply that the process of simplification may be restricted to repeated horizontal interactions between users who are aware of the iconic roots of the signs they use, and who have played an active role in negotiating the simplified form". This, in turn, would mean that "the role of cultural transmission may in fact be severely limited in explaining the emergence of symbols" (Caldwell & Smith 2012: 2). To test whether this is indeed true, instead of dyads, the authors tested multiple chains of ten people each. Notably, though, in each of the seven rounds of the experiment, only four of the ten people who were part of a chain participated. These four participants formed so-called *microsocieties* in which one person was the drawer and three persons served as matchers. Following each round, the drawer, who was simultaneously the most experienced (i.e. 'oldest') member of the group, was replaced by a new, 'naïve' member. To explain this in more detail: in the first round, for example, P1 was the drawer, with P2, P3, and P4 being matchers. In round 2, P1 left the group while P5 was added to it. P2 became the new drawer, with P3, P4, and the newly added P5 being matchers. After that, P2 was replaced by P6, and so on (cf. Table 9). The first six rounds were *transmission phases* in which one of the matchers responded publicly on behalf of all the matchers, sharing with the drawer what they believed a drawing represented. The seventh and final round, in which P10 was added as the last member of the chain, constituted the test round and differed from prior rounds: in this round, all matchers (P8, P9, P10) gave their respective responses privately on a response sheet. Crucially, they had been presented not only the drawings that had been completed by members of their own chain but also drawings by members of other chains.

Table 9. Composition of the test group over rounds of the experiment, from Caldwell & Smith (2012: 3)

| Round number | Drawer | Matchers |
|---|---|---|
| 1 | P1 | P2, P3, P4 |
| 2 | P2 | P3, P4, P5 |
| 3 | P3 | P4, P5, P6 |
| 4 | P4 | P5, P6, P7 |
| 5 | P5 | P6, P7, P8 |
| 6 | P6 | P7, P8, P9 |
| 7 | P7 | P8, P9, P10 |

The clearest result of the study, according to Caldwell & Smith (2012: 7), is that "arbitrary and contrasting conventions were established within microsocieties, and [...] these were transmitted to naïve newcomers to the group who themselves had played no part in negotiating the usage of that particular sign". They report that by round 3, trial durations had decreased significantly compared with durations of the first two rounds, which – together with other factors – indicates that by round 3, drawings had lost their pictographic character and had become arbitrary. One of the most striking findings concerns the apparent advantage of experienced matchers, i.e. matchers who had been participating for at least two rounds, in recognizing drawings from their own microsociety. The authors argue that this advantage does not stem from cues provided by the interactive context, such as drawers' movements during production that could be witnessed by matchers, since these cues were also available to naïve newcomers, who, however, showed no advantage in recognition. Instead, the experienced matchers' recognitional advantage was due to a drawing's similarity to previously produced drawings, i.e. drawings produced in earlier rounds. In the terms introduced in Section 6.2, at this point, *endophoric* (or *relational*) *iconicity* had supplanted exophoric iconicity since what appeared to be relevant at this stage was the "resemblance between sign and sign [and] not between sign and object" (Garrod et al. 2007: 965). In other words, drawers who had been matchers in previous rounds did not attempt to pictographically represent the initial meaning but to establish visual similarity to previous drawings, resulting in endophoric iconicity.

Caldwell & Smith (2012) show even more clearly than Garrod et al. (2007) how pictography can (and is likely to) decrease in writing systems due to *cultural transmission* and *cultural evolution* (cf. also the studies by Fay et al. 2010, Tamariz et al. 2018, and the review in Tamariz 2017). As these terms suggest, the cause for the decrease of pictography is fundamentally sociocultural. It is still mentioned here instead of in the context of the sociocultural fit (cf. Chapter 8) because it is arguably facilitation of production that renders pictography unnatural for production and, in turn, continuous use. The fundamental sociocultural dimension as outlined above makes possible 'only' that this facilitation occurs in an 'orderly' fashion. Conventionalization is possibly a natural reaction to facilitation, ensuring that communication remains possible. If everyone in a community were to simplify shapes in whichever way they wanted to, vast and uncontrolled variation would likely render communication cumbersome, inefficient, or even impossible. The decrease of pictography, initiated by its articulatory unnaturalness, is accompanied by a conventionalization of shapes, which become increasingly arbitrary. Facilitation is a matter of processing (and secondarily communication, since units that are easier to produce are arguably also easier to use in communication), while conventionalization is a sociocultural matter.

What has been discussed so far concerns pictography, i.e. imagic iconicity, and nothing has been said about the effects of diagrammatic iconicity on processing. This is because empirical evidence on the relevance of diagrammaticity is sparse. For language in general, diagrammaticity is claimed to facilitate the acquisition of categories in early language development since it allows exploiting "the regularities in the mapping between representational spaces in different modalities" (cf. Dingemanse et al. 2015: 609). Interestingly, children's productions during early literacy acquisition exhibit features that could be interpreted as diagrammatic: Pontecorvo (1985), for example, describes an Italian 5-year-old who wrote both Italian words for *cane* 'dog' and *cagnolino* 'little dog' using three graphemes, respectively, but the graphemes' basic shapes were bigger when the intended word was 'dog' than when it was 'little dog'. Levin & Korat (1993) mention Israeli children aged 4 to 5 who used more units when they wanted to write 'forest' than when they wanted to produce 'tree', which is highly reminiscent of the diagrammatic Chinese graphemes <木> 'tree' and <森> 'woods' (cf. Section 6.2). Furthermore, in a transmission experiment resembling the ones discussed above, Kirby, Cornish & Smith (2008) found that an artificial language that was culturally transmitted maximized its own transmissibility by 'becoming' more structured, precisely by increasing the degree of diagrammaticity it exhibited.

Findings of this nature are scattered throughout the literature. Randomly mentioning them here does the assumed relevance of diagrammaticity for the processing of writing no justice. Accordingly, a more systematic and elaborate analysis of the interaction between diagrammaticity and processing is necessary for the further development of a functional theory of writing; this will undoubtedly require generating (or locating) more empirical evidence.

### 7.2.3 Indexicality

In their study, Holsanova, Holmberg & Holmqvist (2009) investigated what effect spatial contiguity, the factor most central to indexicality, has on processing. They did so by comparing how readers process different types of complex layouts via an analysis of their eye movements. Research like this is relevant given that most texts that are consumed by readers nowadays are structurally complex. Just consider the typical layout of newspapers: not only different levels of written textual organization – e.g. headlines, captions, quotes, the main body of text – are integral parts of such a complex newspaper layout but also other types of semiotic material including photos, figures, tables, diagrams, etc. It is an obvious assumption that different arrangements of these types of elements affect the reading process in various ways, and this is precisely what is partially captured by the graphematic parameter of indexicality.

Broadly speaking, with respect to indexicality, there are two types of complex texts: first, there are texts that include no explicit connections between text and other elements such as tables or figures. In the reading of such texts, "[w]hen the eyes reach a certain point in the text, it is the reader who has to discover the referential links between the text and the graphics" (Holsanova, Holmberg & Holmqvist 2009: 1216). This is expectedly more challenging cognitively than reading text of the other type in which these links *are* explicitly suggested by features in the text. Such indexical features help guide the reader; they have been studied in the context of the *cognitive theory of multimedia learning* (cf. Mayer 2005). One of these features highlighting indexical relations between textual and non-textual material is captured by the *spatial contiguity principle*: placing elements such as text and an associated figure in physical proximity to one another highlights their relationship, which can render the reading process easier at a cognitive level. Another feature is captured by the so-called *signaling principle*: cues that highlight the organization of a complex text aid reading.

One of the cognitive processes that are influenced by the design of textual layouts is *cognitive load*. According to *cognitive load theory* (cf. Sweller, van Merrienboer & Paas 1998), which, however, focuses on processes of *learning* and not specifically on reading, "instructional materials should be designed to prime the integration of pictorial and verbal representations into a coherent mental representation" (Holsanova, Holmberg & Holmqvist 2009: 1216). Three subtypes of cognitive load make up the total cognitive load: (1) *intrinsic cognitive load* is determined by the complexity of the material and by the learner's competence; (2) *extrinsic cognitive load* is constituted by the design of the learning material – which is where indexicality becomes relevant; finally, (3) *germane cognitive load* is necessary for learners to comprehend the learning material. Studies investigating the effects of indexicality on processing focus on extrinsic cognitive load. The main hypothesis is that explicit indexical cues in a text's layout can reduce extrinsic cognitive load and, in the process, support the other two types of cognitive load by freeing up resources for them.

In their experiment, Holsanova, Holmberg & Holmqvist (2009) tested whether a textual layout that corresponds to the *dual scripting principle* aids the reading process. More specifically, they investigated whether it is beneficial for so-called *attentional guidance*. The dual scripting principle combines the features of spatial contiguity and signaling. The former contributes to an optimal navigation through the text while the latter aids semantic integration, i.e. comprehension, through the conceptual organization of the text. Both principles are relevant to the parameter of indexicality. To test the dual scripting principle's effect, the authors compared readers' eye movements in the reading of infographics that were presented in different layouts. Infographics are a special multimodal text genre that combines illustrations with text and graphic cues such as arrows to describe and explain certain complex phenomena. The two layouts that were presented were (a) a *separated layout* in which illustrations and text were located far from one another and (b) an *integrated layout* in which the elements were spatially contiguous. A second comparison juxtaposed (c) a *radial layout* in which illustrations and texts were not necessarily far apart but arranged in a non-serial way that allowed for multiple entry points and paths for reading with (d) a *serial layout* that was spatially prestructured and (more or less strongly) suggested a reading path. Unsurprisingly, as suggested by readers' eye movements, the integrated and serial layouts were processed more efficiently.

These findings were interpreted partially based on the quantity of *integrative saccades*, defined as "transitions between semantically related pieces of verbal and pictorial information, indicating the process of readers' construction of referential connections between text and illustration" (Holsanova, Holmberg & Holmqvist 2009: 1222; for a more detailed description of eye movements, cf. Section 7.2.8). In the integrated and serial layouts, the mean proportion of integrative saccades was significantly higher than in the separated and radial layouts. Additionally, in the condition radial vs. serial, readers seemed to follow the reading path suggested by the serial layout and spent almost twice as much time reading it, whereas – this is the authors' interpretation – readers quickly became disinterested in the infographic when it was presented in the radial layout since they "have several decisions to make – to choose the entry point, to decide about the reading path, to find relevant pieces of information, to create a connection between them and to integrated [sic] them mentally" (Holsanova, Holmberg & Holmqvist 2009: 1224).

A single study evidently does not suffice to proclaim the most natural configurations on the parameter of indexicality when it comes to processing. It does, however, point to the relevance of indexical phenomena in the reading of texts. Notably, the fact that indexicality, as conceptualized here, is a universal feature of text design that is largely independent of the writing system (and script) in which said design is materialized means that this parameter is also not particularly useful in comparing different writing systems or scripts. This, too, warrants its only brief treatment in this chapter. In fact, indexicality is, although it is certainly central in an analysis of the written modality of language, a parameter that is more fittingly analyzed within the theoretical framework of Natural Textlinguistics in which it was described in some detail (cf. Dressler 1989: 23–35). Indeed, Natural Textlinguistics focuses on the naturalness of (mostly written) texts and textual devices and does so predominantly from an intra-textual perspective. This is precisely the scope of the effects of indexicality.

### 7.2.4 Transparency

The questions of if and how transparency affects processing are central, partially because transparency is a parameter with respect to which even writing systems of the same type such as alphabets vary remarkably: it suffices to consider the vast differences between the transparencies of English and Finnish, for example. Thus, parameters such as transparency can provide a basis for more fine-grained comparisons of diverse writing systems that go further than the established (broad) typological classifications. This section discusses the question of how transparency affects processing. Here, the focus is on perception since transparency, which by definition is a written unit's property of consistently corresponding with only a single linguistic unit, plays a more crucial role in reading. By contrast, the inverse parameter of uniformity (see below) takes center stage in production, i.e. writing/spelling.

No discussion of transparency can do without mentioning the so-called *Orthographic Depth Hypothesis* (ODH). This well-known hypothesis postulates that more transparent writing systems should be easier to read as they allow more direct access to phonology (cf. Katz & Frost 1992). This connection is based on the obvious assumption that reading, or more specifically, those word recognition processes that rely on phonology, are facilitated by transparent semiotic relations between basic shapes and phonemes. As this characterization implies, the ODH is concerned with phonographic transparency and in its original form can capture only effects pertaining to the processing of phonographic writing systems or phonographic elements in writing systems (cf. Daniels & Share 2018). Thus, the ODH would predict, for instance, that reading morphographic kanji in Japanese is harder than reading syllabographic kana since kanji allow no access to the morphemes' phonological representations, which is the central criterion for the ODH.

Such a one-dimensional and phonocentric interpretation of transparency obviously disregards other types such as morphographic transparency. When a morphographic grapheme corresponds with a morpheme in a transparent manner, this arguably also has beneficial effects for processing regardless of whether the grapheme offers phonological information. While the ODH places its emphasis on phonographic aspects of the reading process, in its weak form, it acknowledges that phonology is not the sole variable relevant for reading (cf. Katz & Frost 1992: 72). Although morphography plays a role not only in morphographic writing systems but, as a 'secondary' principle, also in phonographic systems (cf. Schmidt 2018 for German), morphographic transparency has not yet gained currency in reading research (at least not as much as phonographic transparency). Thus, there is a lack of studies tackling the question of how the (in)transparent representation of morphemes or morphological information affects processing, and this is the case even for morphographic writing systems (see below).

By contrast, there is an abundance of studies that provide evidence that phonographic transparency is a relevant parameter for processing (especially perception, i.e. reading). Thus, phonographically transparent writing systems are acquired faster (cf. Seymour, Aro & Erskine 2003) and are less error-prone than non-transparent systems (cf. Cossu et al. 1995). Furthermore, they are read aloud more quickly by children than non-transparent writing systems, which is interpreted as an indication for generally better literacy development (cf. Ellis et al. 2004). Additionally, in conceptions that assume a reciprocal relationship between phonological awareness and reading, phonographic transparency is claimed to benefit the development of phonological awareness on the basis that "one-to-one mapping between letters and sounds in a transparent orthography promotes access to phonemes, thus boosting basic phonological-awareness skills and helping to trigger the development of phoneme-sized representations" (Ziegler et al. 2010: 556; cf. also Castles & Coltheart 2004).

The literature on the effects of phonographic transparency (often referred to as *orthographic transparency*[250]) is massive, and it is not my goal to pro-

---

**250** Once again, the question of whether this type of transparency is graphematic or instead 'orthographic', as it has been called in the literature, is justified. When reading, readers of most writing systems are in fact confronted with orthography, i.e. 'correct' writing. However, since transparency is (unlike, for example, compositional transparency) a paradigmatic parameter (cf. Section 6.4), it is concerned with segmental relations between basic shapes and linguistic units, and these, I argue, are not reguvide a comprehensive picture of it here. The evidence cited above suggests strongly that the transparency of the graphematic module that was assessed descriptively is linked not only to a natural linguistic fit but also to a natural processing fit. This is by no means a groundbreaking conclusion, but it leaves open the important question of the conflict between phonographic and morphographic transparency. Both can be attained simultaneously only in 'phonologically simple' languages such as Serbian and Croatian. In cases in which these two types of transparencies converge, "morphologically related words have a common phonologically invariant core" (cf. Katz & Frost 1992: 70).

Commonly, however, there is a tradeoff between them, meaning phonographically transparent systems are morphographically intransparent and morphographically transparent systems are phonographically intransparent. In this context, an observation that is made frequently is that beginning readers tend to rely on phonographic transparency, whereas skilled readers rely more on morphographic transparency (cf. Daniels & Share 2018: 108). In German, for example, morphography is explicitly regarded as a graphematic aid for reading (cf. Fuhrhop & Schreiber 2016): morphographically transparent consonant doubling (such as <nn> as in <können> 'can') has a beneficial effect on processing (cf. Bredel, Noack & Plag 2013). The gradual transition from phonological decoding to morphological decoding happens not only in phonographic systems but also in morphographic ones: children learning to read Chinese first focus more on the phonological components of graphemes when they read but gradually come to rely more on cues of meaning provided by the semantic components (cf. Tong, Tong & McBride 2017). Ontogeny is echoed in phylogeny, as there is a trend for writing systems to develop from being phonographically transparent to becoming lexically distinctive, i.e. morphographically transparent, which Sampson (2018) demonstrates for English. Note, again, that in phonologically simple systems (i.e. those lacking complex morphonology), these two types of transparency are not in conflict in the first place.

In sum, it is obvious that transparency is a parameter central for the graphematic processing fit as it not only makes possible perceptual decoding processes but also renders them more efficient. Which subtype of transparency is more central, however, meaning which linguistic level (or type of linguistic information) is represented transparently by the graphematic module, is a question that must be assessed individually for each writing system and independently of its type. It is ultimately a question that concerns the linguistic fit as the structure of a given language mandates which type of transparency is more important. Given the above-mentioned findings on phonographic transparency, transparency in general is expected to influence processing in fundamental ways. However, more

lated by the orthographic module, which regulates sequences of graphemes (mainly written words) instead. Whether /s/ is spelled <s> or <ß> in German, for example, cannot be decided segmentally but only syntagmatically within a larger context. In such cases, both spellings are located inside the graphematic solution space of a word but only one of them is orthographically correct.

studies are needed on how users benefit from morphographic transparency in phonographic systems and phonographic transparency in morphographic systems.

### 7.2.5 Uniformity

Because uniformity implies the analytical direction of *language → writing*, it is expected to affect spelling processes more than reading processes. As established in the context of the linguistic fit (cf. Section 6.5), writing systems whose graphematic modules exhibit a high degree of transparency can be but are not necessarily uniform, and vice versa. Unsurprisingly, in transparent but non-uniform writing systems, spelling is more difficult than reading (cf. Bosman & Van Orden 1997) because writers must choose between multiple possibilities of writing a given linguistic unit that are all located inside the graphematic solution space (cf. Geva, Wade-Woolley & Shany 1993: 384). Consequently, spelling acquisition often lags behind reading acquisition as it is "a challenge in spelling [...] to specify orthographic units that may be irrelevant for reading" (Gingras & Sénéchal 2019: 37). This has been shown for a number of systems, including Thai (cf. Winskel 2014), German (cf. Wimmer & Mayringer 2002), French (cf. Sprenger-Charolles, Siegel & Bonnet 1998), Hebrew (cf. Geva, Wade-Woolley & Shany 1993), Brazilian Portuguese (cf. Pinheiro 1995), and Persian (cf. Rahbari, Sénéchal & Arab-Moghaddam 2007), and the interested reader is referred to the pertinent studies for details. In a (very brief) nutshell, non-uniform graphematic relations between linguistic units (in all above-mentioned systems, these are phonemes) and basic shapes result in difficulties in production.

### 7.2.6 Compositional transparency

As already lamented in the context of the linguistic fit, graphematic representation of vowel phonemes in Thai is often complex (cf. Section 6.6). Concerning the question of how this affects processing, an experimental study by Winskel (2010: 26) shows that the spelling of complex vowels is more error-prone and acquired later by children. She attributes this to 'memory load', which she claims is higher for the spelling of complex vowels than for simple vowels.

In Arabic, compositional opacity, specifically the fact that short vowel phonemes remain unrepresented, affects the processing of graphematic words. Usually, the compositional value of a graphematic word, e.g. <CCC>, does not correspond with its phonological representation, e.g. /CVCVCV/. This is an example of graphematic underspecification. However, as short vowels *can* be written, experiments have tested whether a transparent representation of *all* phonemes is more natural for processing. Notably, existing studies have yielded inconclusive results with respect to this question. For example, in an investigation of accuracy and fluency in reading aloud, Asadi (2017) tested 1516 Arabic-reading children from the first to sixth grades. Results indicate that unvowelized Arabic is read more accurately and fluently in all grades except for the first and second grades, where reading accuracies are similar for the vowelized and unvowelized conditions. In general, unvowelized Arabic appears to be more natural for reading than vowelized Arabic (cf. Saiegh-Haddad & Schiff 2016; Taha 2016). There are at least two possible explanations for these findings: firstly, vowelization makes Arabic graphetically more complex, as it increases the visual material that needs to be processed and contributes to "a visual load that may interrupt fluency" (Asadi 2017: 139). Secondly, while there are possible benefits of vowelization for reading, these are a function of grade level (cf. Shany, Bar-On & Katzir 2012 for similar results for Hebrew). This has several reasons, one of which is that at the beginning of literacy instruction, children are taught the vowelized version of the writing system, while the unvowelized variant is introduced later, in the third and fourth grades (cf. Asadi 2017: 138). A second reason is that beginning readers appear to be more reliant on transparent phonographic information since the 'orthographic lexicon' is not yet developed and direct access to meaning through reading is not (yet) as fast as through phonological access (cf. Asadi 2017: 144; Taha 2016: 139f.).

However, as implied above, the findings of several studies contradict the assumption that full phonographic compositional transparency is beneficial. This, as already mentioned, is claimed to be due to visual complexity (cf. Ibrahim 2013; Taha 2016; Asadi 2017). By contrast, other studies (e.g. Abu Rabia 1998, 2001) have pointed to a facilitating effect of vowelization on reading accuracy and comprehension. Additionally, there is evidence that vowelization helps disambiguation of heterophonic homographs, i.e. graphematic words that are written the same way in their unvowelized versions but have different phonological representations. By contrast, vowelization has no effect on unambiguous words (cf. Maroun & Hanley 2017).

A different type of compositional opacity, graphematic excess, is exhibited, for example, by graphematic words that include 'silent' graphemes, i.e. graphemes "that do not serve a phonological function" (Gingras & Sénéchal 2019: 37). 'Silent', of course, is a phonocentric term stemming from the fact that silent graphemes have been described predominantly for phonographic writing systems.[251] A recent study of silent graphemes in French showed that it requires more cognitive effort to process silent graphemes than graphemes that encode phonological information. A possible explanation is that "nonphonological letters are more likely to be omitted from orthographic representations given their lack of phonological value" (Gingras & Sénéchal 2019: 44f.). Descriptively, this can be captured by the fact that a silent grapheme is not associated with a segmental signatum the way graphemes that relate to phonemes are. As such, it does not contribute to the phonological representation of a graphematic word. Note, however,

251 Graphematic excess is not restricted to phonography. Morphographic systems, too, can include graphematic information that does not represent any linguistic information.

that in some cases, 'silent' graphemes may serve other functions, such as the plural morpheme <-s> in <amis> 'friends'. The grapheme might not be relevant phonologically (at least in isolation)[252] but it is relevant morphologically. Notably, however, it is acquired by children much later than phonographic French graphemes (cf. Totereau, Thevenin & Fayol 1997).

### 7.2.7 Positional transparency

Since 2003, a text referring to (fictitious, cf. Velan & Frost 2007: 913) research supposedly conducted at Cambridge University has circled the internet. It is characterized by many words whose internal graphemes appear transposed while their initial and final graphemes are kept intact. It reads:

> Aoccdrnig to a rscheearch at Cmabrigde Uinervtisy, it deosn't mttaer in waht oredr the ltteers in a wrod are, the olny iprmoetnt tihng is taht the frist and lsat ltteer be at the rghit pclae. The rset can be a toatl mses and you can sitll raed it wouthit porbelm. Tihs is bcuseae the huamn mnid deos not raed ervey lteter by istlef, but the wrod as a wlohe.[253]

The ongoing (and to a large degree public) fascination for this text is based on the fact that although the graphemes are 'jumbled', the text, as it itself points out, remains readable and understandable. This effect was termed – based on the alleged research referred to in the text – *Cambridge University effect* and has attracted the attention of the reading research community. Notably, it poses severe problems for models of word recognition such as the *IA model* (cf. McClelland & Rumelhart 1981) and the *dual-route cascaded model* (cf. Coltheart et al. 2001), which are position-specific in that they code grapheme positions explicitly and absolutely (cf. Section 7.1.3). The awareness of these position-based shortcomings of existing models spawned several new models of reading (particularly word recognition) that allow for varying degrees of flexibility concerning grapheme positions. They include the *SERIOL model* (cf. Whitney 2001), the *SOLAR model* (cf. Davis 2010), the *open bigram model* (cf. Grainger & Van Heuven 2004), and the *overlap model* (cf. Gomez, Ratcliff & Perea 2008).

With respect to the linguistic fit, positional transparency was defined as the alignment of corresponding elements in the signans and the signatum. In other words, that a series of graphemes <abc> corresponds with a sequence of phonemes (or syllables or morphemes) /abc/ (cf. Section 6.7). Whether this is relevant in processing can best be evaluated by investigating how readers process positional mismatches. It is important to note that these 'mismatches' are not necessar-

252 When larger contexts are considered, silent graphemes such as the <-s> in <amis> can become phonographic due to liaison.

253 This is taken from https://www.mrc-cbu.cam.ac.uk/people/matt.davis/cmabridge/ (July 22nd, 2020), where Matt Davis from Cambridge University dissects the text and its claims and offers explanations and useful references.

ily the result of errors that lead to graphematically illegal spellings but, in some systems, idiosyncratic features. The most prominent example of this has already been mentioned multiple times: Thai and its misaligned vowel graphemes. In the course of four reading and spelling experiments, Winskel (2009) examined how these graphemes are processed by Thai adults and children. The results suggest that there is no significant processing cost associated with those misaligned vowels that operate *within* syllables (which Winskel calls Type 1), whereas misaligned vowels that operate *across* syllable boundaries (Type 2) are a frequent source of errors, especially for children during literacy acquisition.

An example of a Type 1 word is เพลง <ephlŋ> /phleŋ/ 'song', an example of Type 2 แมลง <ɛːm(a)lŋ> /m(a)lɛːŋ/ 'insect'. In the Type 1 word, the vowel <เ> = <e> is written in initial position although in the corresponding phonological representation, it follows the consonant cluster <พล> = <pl>. Given that the word is monosyllabic, the scope of this misalignment is the syllable, it is intrasyllabic. By contrast, in the Type 2 word, the vowel <แ> = <ɛː> is written word-initially, i.e. at the beginning of the first of two syllables despite being the medial vowel of the second syllable. The first syllable is <ม> = <m(a)> (with (a) indicating an implicit vowel). The misaligned vowel <ɛː> fits into the second syllable between the consonants <ลง> = <lŋ>. These Type 2 words in which misaligned vowels are positioned across syllable boundaries produced the greatest processing cost for Winskel's participants. This provides additional evidence for the above-discussed claim that the syllable is the preferred unit of processing (cf. Section 7.2.1). In the case of Thai, the syllable is not the unit of representation, since graphemes represent phonemes (or, sometimes, sequences of phonemes) rather than syllables. However, the syllable still appears to be the relevant psychological grain size for reading in Thai (cf. Ziegler & Goswami 2005). In other words, it is the unit of processing, as Winskel (2009: 19) also suggests in the conclusion of her study: it is "a more natural segmentation unit than the phoneme in reading Thai in adults and children". This brings us back to the question of which concept of 'syllable' is meant here: the phonological syllable or an independent graphematic syllable (cf. Section 7.2.1)?

In this vein, evidence from Japanese suggests that positional transparency is "orthographic in nature", i.e. independent of phonology (Perea, Nakatani & van Leeuwen 2011: 700). In a silent reading experiment, participants were presented target words in syllabographic kana with different following words in the parafoveal region, i.e. the region that is not in the focus of the eyes' fixations but from which some visual information can be extracted (see also below). In the experiment, the words located in the parafoveal region either included a transposed syllable (<アリメカ> <a.ri.me.ka.> instead of <アメリカ> <a.me.ri.ka>, with the transposed syllables underlined) or replacement syllables (<アカホカ> <a.ka.ho.ka> instead of <アメリカ> <a.me.ri.ka>). Results showed that fixations on the target words preceding the manipulated parafoveal word were shorter when the parafoveal word included a transposed syllable than when it included a replacement syllable. This confirms the *transposed letter effect* (or, in line with the more inclusive terminology proposed here, *transposed grapheme effect*). In general, this effect predicts that a prior presentation of, for example, *<jugde> facilitates processing of a word such as <judge>. By contrast, processing is not facilitated when the prime is, for instance, <jukpe>, and thus includes different instead of merely transposed graphemes.

A central question is at which linguistic level positional transparency is relevant for processing. In other words: what is its scope? Aside from the syllable, the morpheme and the lexeme have been discussed as relevant domains. In this context, Christianson, Johnson & Rayner (2005) found that English primes that contain transposed graphemes morpheme-internally elicited a transposed grapheme effect while in primes in which transposition occurred across morpheme boundaries, no transposed grapheme effect was observed. This is in line with the findings of Duñabeitia, Perea & Carreiras (2007) who showed that both in agglutinative Basque and non-agglutinative Spanish, transposed grapheme effects occurred intra-morphemically in non-affixed words, whereas no such effects occurred in affixed words across morpheme boundaries. In yet another study investigating transposition effects in Basque, Perea & Carreiras (2006: 421) concluded that transposition effects were not affected by lexeme boundaries in compounds on the basis that transposition effects were very similar when they occurred in non-compound words to when they occurred across lexeme boundaries in compound words. The findings of these studies are not contradictory, as Duñabeitia, Perea & Carreiras (2007) examined morpheme boundaries between lexemes and grammatical affixes while Perea & Carreiras (2006) investigated morpheme boundaries between lexemes, i.e. lexeme boundaries. These different types of boundaries appear to differ with respect to whether they are relevant for transposition effects – at least in Basque.

Another important question is whether the position of graphemes that are jumbled is of relevance for processing. This appears to be the case in writing systems using Roman script, for example, where transposition of the initial grapheme (e.g. *<rpoblem> instead of <problem>) challenges visual word recognition to a larger degree than transposition of internal graphemes (e.g. *<porblem>), suggesting that the initial grapheme enjoys a privileged status (cf. White et al. 2008), which explains why the 'Cambridge university effect' (see above) works so well. By contrast, in their eye movement study, Winskel, Perea & Ratitamkul (2012) found that in Thai, the initial grapheme of a word does not have such a privileged status. While in Thai, silent reading of sentences that include nonwords with transposed graphemes proves more difficult than reading sentences without such grapheme transpositions (as is expected), there appears to be no difference in difficulty between transposed initial and transposed internal graphemes. This implies that the "actual identity of the letter is more critical than letter position in Thai" (Winskel, Perea & Ratitamkul 2012: 1532). The authors further theorize that this is because processing is modulated by the specific characteristics of the writing system. That the writing system of Thai, as detailed above, exhibits misaligned vowels as an important graphematic feature appears to perpetuate a certain degree of positional flexibility for processing written Thai. This highlights the relevance of the system-specific level of naturalness. Positional transparency is still arguably most natural at a universal level, but the degrees of how much violations of positional transparency disturb processing vary across writing systems depending on their specific characteristics.

Arabic and Hebrew offer additional striking examples of the relevance of system-specific naturalness. In both systems, only consonants and long vowels are graphematically represented. The consonant graphemes highlight the (tri)consonantal roots that occupy a central role in Arabic and Hebrew morphology. There is only a minimal (if any) degree of flexibility when it comes to the position of these consonant graphemes within a word. An example illustrates how transposition changes meaning: the Hebrew root S.L.X means ‘to send’, while X.L.S means ‘to dominate’, X.S.L means ‘to toughen’, and L.X.S means ‘to whisper’ (cf. Velan & Frost 2007: 914). Evidently, consonant graphemes that make up roots cannot be transposed without causing severe difficulties in recognizing the intended (as opposed to a different) word. In a study, it was found that grapheme transposition in Arabic words in which the order of the root graphemes was kept intact produced priming effects, whereas words in which root graphemes were transposed did not because the latter can have “the negative impact of activating the ‘wrong’ root family” (Perea, Mallouh & Carreiras 2010: 378). The authors additionally emphasize that (similar to what was found for Thai) there are no significant differences between initial and internal grapheme transpositions, as transposition effects are similar regardless of grapheme position as long as the sequence of root graphemes is kept intact.

ئىنتايىن inta_jin
ئىتنايىن *itna_jin
سوۋغات so_w_ʁa_t
سوغۋات *so_ʁw_a_t

Figure 47. Transposed graphemes with changing ligations

Another system-specific feature of Arabic that interacts with positional transparency is positional allography (cf. Section 2.3). In this vein, Yakup et al. (2014) tested whether transposition effects differ for words in which grapheme transposition leads to the occurrence of a different basic shape (because it relocates the grapheme to a different type of position) as opposed to words in which the basic shape remains the same after transposition. Notably, if the basic shape changes, so do the ligations, i.e. the connections to preceding and/or following graphemes, which marks a second important variable. The language used in the experiment was Uyghur, an agglutinative language spoken in China that uses Arabic script but, unlike other writing systems using Arabic script, graphematically represents all vowels (including short vowels). An example of a word with transposed graphemes in which both basic shapes and ligations were retained (i.e. stayed the same) is *<itna_jin> (from the correct word <inta_jin> ‘very’). In these words, the two transposed graphemes <n> and <t> are both embodied by the medial allographs both in the correct and transposed conditions. The underline in the examples indicates where these words exhibit empty spaces determined by the graphemes that remain unconnected to their left. An example of a pair in which the transposition caused a change in ligations is *<so_ʁw_a_t> from correct <so_w_ʁa_t> ‘gift’ (cf. Figure 47), where <w> no longer appeared in its isolated form but in the final form that is connected to the right. The study’s results show that Uyghur readers had more difficulty reporting target words when grapheme transposition involved changes in positional allographs and ligations than when it caused no such changes (cf. Yakup et al. 2014: 1604). Interestingly, the fact that grapheme transposition involving different basic shapes appears to hinder transposition priming effects is interpreted by the authors as evidence that grapheme position is not encoded at the level of graphemes (‘abstract grapheme identity’) but at the level of basic shapes, specifically of positional allographs.

With respect to the latter point, the study of Friedmann & Haddad-Hanna (2012) is worth mentioning. They tested Arabic-speaking individuals with letter position dyslexia (LPD), which is “caused by a selective deficit to letter position encoding […] which results in migration of letters within words, primarily of middle letters” (Friedmann & Haddad-Hanna 2012: 193). Thus, in this case, transposition is not present in the graphematic input but is a pathological process that occurs during processing; terminologically, this is underlined by the fact that it is referred to as ‘migration’. Subjects were presented with ‘migratable words’ in which the migration of graphemes could either cause (1) a ‘form change’, i.e. a change of positional allographs in medial graphemes, (2) a change of the final grapheme, (3) a change in ligations without an accompanying change in form, or (4) no change at all. Participants made fewer migration-induced errors in reading words out loud when grapheme migration would have caused a change in allographs than when it would not have (cf. Friedmann & Haddad-Hanna 2012: 197). Accordingly, positional allographs appear to be helpful for individuals with LPD. This, in turn, points to the possibility that positional information is stored in the ‘abstract grapheme identity’, i.e. the abstract representation of a grapheme. If this were true, there would be a discrepancy between grapholinguistic description and cognitive reality: descriptively (at least in the approach presented here), positional allographs are assigned to the same abstract grapheme. However, in processing, they appear to be stored independently, which would mean that cognitively, they have graphematic status, i.e. are treated and stored as separate graphemes.

At the end of this section, it is important to note that positional transparency can not only be assessed for grapheme sequences but also at the subsegmental level, i.e. concerning the spatial arrangement of components within a grapheme. Subsegmental components in Chinese graphemes, for example, occur in a preferred position depending on their function: semantic components occur on the left, phonological components on the right (cf. Taft, Zhu & Peng 1999: 498; Ho, Ng & Ng 2003: 851). And indeed, among the graphemes consisting of a semantic and a phonological component, 75% have their semantic component on the left (cf. Feldman & Siok 1999). This preferred position of components is "the crux of a character's orthographic structure" as it helps to "determine whether the character is legal or not" (Ho, Ng & Ng 2003: 853). If components are in their legal positions, the grapheme in question is either an existing grapheme of Chinese or a legal pseudographeme, i.e. a grapheme that conforms to the systematics of the system but does not actually exist in the system (which makes it a systematic gap, cf. Section 5.1). By contrast, if any component appears in the illegal position, the result is a nongrapheme. Knowledge of component position is predominantly implicit and is tested in character decision tasks in which participants are asked to judge whether a grapheme is legal (i.e. an existing grapheme or a pseudographeme) or not. Even first graders can make use of this information and reject nongraphemes while judging pseudographemes as acceptable (cf. Shu & Anderson 1999). The question of whether the representation of components is position-invariant or position-specific has not been conclusively answered; some scholars posit invariance (cf. Taft, Zhu & Peng 1999), i.e. that there is an abstract representation subsuming all positional allographs. However, recent evidence gained through ERPs suggests that the representation of components might be position-specific (cf. Su et al. 2012).

### 7.2.8 Figure—ground

One of the perceptually most salient features of writing is the fact that written marks stand out visually against their background. This is achieved by means of the spaces of the writing surface that shine through between (and within, as in |O|) the shapes that make up scripts. In the gestalt theoretical terms that were borrowed by Natural Morphology, the writing surface is the ground, the shapes are the figures. Some of these empty spaces on the writing surface make visible graphetic units of various sizes (cf. Section 1.2). These units then function as different graphematic units across writing systems. Notably, the size of graphetic chunks and, in turn, graphematic units that are separated visually by empty spaces differs across the world's writing systems.

Spacing, as the most relevant manifestation of a figure—ground distinction, is located at the interface between graphetics and graphematics. Like figure—ground in general, spacing is largely a perceptual parameter; as such, it influences the legibility of text. However, since spacing does not function randomly but in most cases visualizes graphematic, i.e. linguistic units, it is also a central graphematic parameter. Accordingly, it renders texts not only more legible but also more readable. This hybrid functionality of spacing is reflected in the ways it affects processing: as a graphetic tool, it aids the guidance of eye movements in reading while as a graphematic tool, it facilitates the recognition of linguistic units such as words; the latter helps readers extract linguistic information and, in turn, guides parsing processes in which syntactic structures are constructed. In the following, the focus will be on how spacing affects eye movements, which shall first be characterized briefly.

At the beginning of the reading process, light waves impinge on the eye. There, they are perceived by light-sensitive cells, so-called photoreceptors, and project two slightly different images onto the right and left retina. This is followed by sensory transduction, in the course of which the energy of the sensorial stimulus is converted into electric signals; these are then transferred to the brain (cf. O'Shea 2012: 17). Reading, ultimately, is a complex bundle of processes: the physiological part of the task is performed by the eyes, while the cognitive task is performed by the brain. In this respect, Smith (1971: 82) cautions that "[r]eading should not be regarded primarily as a visual process" as the brain not only guides eye movements and regulates when the eyes can take in visual information but also determines which and how much information is taken in; this is important given that the eyes can (and do) take in more information than the brain can process. Reading, thus, is fundamentally a selective process in which the brain selects which parts of the visual stimulus are important. And as long as the brain is processing this information, the eyes cannot take in new information and send it to the brain.

In fact, since the brain is constantly processing information while reading, the eyes can only make roughly four 'snapshots' of what they perceive per second (cf. Smith 1971: 90). Considering this basic architecture of the reading process, two aspects of what our eyes do prove central: *when* the eyes move and *where* they move to. The eyes move in the direction of reading (whether that is left to right, right to left or top to bottom) but do not do so in a linear fashion; also, they do not stop to fixate every word. Rather, they 'jump', and these jumpy movements are called *saccades*. Between saccadic movements, the eyes stop at certain points that they fixate. Fittingly, these breaks in the movement of the eyes are termed *fixations*. Importantly, it is only during these fixations that the brain makes (or tells the eyes to make) the above-mentioned 'snapshots' and processes the information perceived in them (cf. O'Shea 2012: 13f.). These fixations usually last between 200 and 250 milliseconds; the spatial scope that is perceived in them spans roughly 8 segmental spaces (cf. Rayner 1979: 61; O'Shea 2012: 14; Gibson & Levin 1980: 195), although both the duration and the visual intake of fixations vary. The totality of the perceptual span (including the parafoveal and peripheral regions, see below) amounts to four basic shapes to the left and 15 basic shapes to the right of the point of fixation, at least in writing systems that use Roman script (cf. Rayner, Well & Pollatsek 1980). At the graphematic level, in alphabets, grapheme identity can be extracted from roughly 6–8 graphemes to the right of the fixation point (cf. Underwood & McConkie 1985).

Sometimes, the eyes do not move forward in the text, but backward; these movements are referred to as *regressions*. While their exact function is still a matter of discussion, it is clear they allow rereading and refixating words, e.g. because they were not read correctly the first time (cf. Booth & Weger 2013). Regressions are also common in the reading of so-called *garden path sentences* for which the first parsing is incorrect (which readers realize at some point during the reading of the sentence); this requires readers to go back and re-parse the sentence (cf. Altmann, Garnham & Dennis 1992). Broadly speaking, regressions are a reflection of heightened cognitive activity.

What proves essential for an analysis of the relevance of spacing is a categorization of different regions of visual information that is perceived by the eyes during the movements they make. Rayner (1979: 61) distinguishes between the *foveal, parafoveal*, and *peripheral* regions. The foveal region is what is in the eyes' focus during a fixation: vision in the foveal region is sharp and the information located in it is perceived clearly. By contrast, while the parafoveal region is not in focus, information located in it can still be perceived and, to some degree, taken in. This is crucial for understanding the relevance of spacing: as will be shown below, spaces located in the parafoveal region are relevant in saccade programming, i.e. for the eyes and the brain to decide on the target of the next saccade, which is the location of the next fixation. In experiments in which readers were deprived of parafoveal input by masking words located to the right of the currently fixated word, it was found that reading rates dropped to 60% of reading rates that are expected in the normal condition in which the whole line is visible and extraction of parafoveal information is possible. The fact that information obtained from the parafoveal region guides eye movements is referred to as *parafoveal preview benefit* (cf. Rayner 1998). Finally, in the peripheral region, which is located to the outermost of the currently fixated region, movements and light contrasts can be perceived.

In the comparison of writing systems, one of the most ostensible differences is whether they exhibit spacing between 'words' or not (cf. Section 2.5). In general, spaces between words guide eye movements during the (silent) reading process as characterized above. Specifically, they offer parafoveal cues that help in programming saccades, which – in alphabets, particularly English – are usually targeted between the beginning and the center of the word that shall be fixated; this specific position is referred to as the *Optimal Viewing Position* (OVP, cf. O'Regan 1990) or *Preferred Viewing Position* (PVL, cf. Rayner 1979). Unsurprisingly, removing word spaces from writing systems of which they are an integral feature is detrimental to reading: for example, reading processes in users of alphabets employing Roman script are slowed down by 30–50% when word spacing is removed, which is a result of the disruption of eye movement control (especially parafoveal processing) and word identification (cf. Morris, Rayner & Pollatsek 1990; Sheridan, Reichle & Reingold 2016). Adults' eye movements while reading unspaced text resemble eye movements in children who have not yet mastered silent reading (cf. Fisher 1976). If empty spaces are missing, the beginnings and endings of words are concealed, information that is crucial for word identification. In terms of eye movements, removing empty spaces interferes with the programming of saccades as vital spacing information is removed from the parafoveal region. As a result, saccades are shorter and land closer to the beginning of words than to the above-mentioned OVP; at the same time, fixation durations become longer (cf. Rayner, Fisher & Pollatsek 1998; Perea & Acha 2009: 1994).

Information about word boundaries does not necessarily come from spaces located between words, as Perea & Acha (2009) have shown: alternative forms of indicating word boundaries such as the alternating**bold** manipulation (unspaced text in which each other word is printed in bold) come at "some reading cost" (Perea & Acha 2009: 1999) compared to normal spaced text but leave word identification and guiding of eye movements mostly unhindered. The alternating**bold** scenario is comparable to cues provided in the writing system of Japanese (see below) in which it is not word spacing that provides clues about word boundaries but the alternation of the component scripts (and their different functions).

Chinese, Japanese, and Thai (also Lao, Khmer, Balinese, and Tibetan, cf. Winskel 2016: 152) are often named as examples of writing systems that lack spacing between words. For this reason, they are often used in experiments that aim to study the effect of word spacing on reading, specifically the question whether it benefits reading in systems that lack it; this would be the opposite effect of the disruption of reading due to the removal of spaces from systems that usually exhibit them. Some of the relevant research shall be mentioned in the following to evaluate the status of this central manifestation of figure—ground for the graphematic processing fit.

For Thai, which has no empty spaces between words but rather between syntactic units, various studies have tested whether the introduction of word spacing facilitates reading. An alternative hypothesis is that it disrupts reading because it introduces a feature that experienced readers of Thai are unfamiliar with (cf. Winskel 2016: 153). These two diametrical effects underline the tension between the levels of universal and system-specific naturalness: while word spacing is assumed to be natural for processing universally, the fact that some users have acquired literacy in writing systems lacking this feature might cause a shift in what is natural for them in processing. One of the first studies conducted to investigate this question found that while spacing in Thai did not facilitate targeting of saccades or early lexical segmentation, it did help later word processing (cf. Winskel, Radach & Luksaneeyanawin 2009). First fixation durations (i.e. durations of the first fixation of a word, which excludes fixations made after regressions) did not differ in the spaced and unspaced conditions and the same applied to the default fixation landing position, which was to the left of the word center in both conditions (cf. Reilly et al. 2005). These findings were replicated in a later study (cf. Winskel, Perea & Ratitamkul 2012), suggesting that word spacing affects eye movements in Thai neither positively nor negatively. However, it was found to facilitate word processing, i.e. comprehension (cf. Winskel 2016: 158).

The Japanese writing system is a compelling case for experiments on word spacing since the systematic alternation of the three component scripts (kanji and the two kana scripts) already provides information about where word boundaries are located (cf. Section 2.5). Indeed, when the system is tested in its default mixed form, i.e. with the alternation of scripts, the insertion of word spacing leads to similar effects as in Thai. Spaced text is read slightly more slowly than normal unspaced text, although not significantly (cf. Sainio et al. 2007). Interestingly, when text was presented exclusively in kana (hiragana, to be specific), the results more closely resembled the results obtained in spacing experiments that tested English: spacing in Japanese kana aided reading, although to a lesser degree (12%) than in English (30–50%). Specifically, word spacing in kana texts had a facilitative effect on eye movements and word identification. This is not surprising since the substitution of kanji by kana rids text of the word boundary cues usually provided by the alternation of these scripts. Consequently, segmentation must proceed in a way that Japanese readers are not used to (at least from their native system; they are of course familiar with it when they have been literalized also in an alphabet). Against this background, it is interesting to note that in Japanese literacy instruction, children are first exclusively taught the hiragana syllabary. Since texts presented in only hiragana offer no word boundary cues, oral group recitation of texts forms a crucial part of elementary reading instruction (cf. Sakamoto & Makita 1973). This echoes the practice of reading aloud in order to segment words in stages of Latin and Greek literacy in which word spacing was not yet available (see below).

In Chinese, the introduction of word spacing was also experimentally tested. Chinese differs from the just mentioned Japanese in that it uses only one script, meaning there is no script alternation and no associated visual information about where word boundaries are located. The fact that Chinese has been (and to some degree still is) commonly called a 'logographic' writing system should not conceal the fact that it is in fact a morphosyllabographic system: graphemes relate to morphemes (instead of words, which Greek *lógos* refers to), and these are mostly monosyllabic. Now, the fact that in contemporary Chinese, 66% of words are bisyllabic (that is, 66% of 55,000 unique words in the *Academia Sinica* corpus), means that 66% of written words are made up of two graphemes and, in turn, two basic shapes (cf. Yen et al. 2012: 1009). This prevalence of bisyllabic Chinese words (as well as trisyllabic and tetrasyllabic words, which are less frequent than bisyllabic words but more frequent than monosyllabic words) renders the absence of word spacing in some ways a challenge for the reading process because it means words are largely not made up of single graphemes which are separated from each other by spaces. To make matters worse, in an analysis of the *Academia Sinica* balanced corpus, Yen et al. (2009) found that only 17,8% of 5,915 unique graphemes unambiguously signal word boundaries; moreover, the frequency with which these boundary graphemes occur is low. Thus, word boundaries are not reliably indicated by them. In stark contrast, 49% of the 5,915 tested graphemes can be used in every position in a word. This positional ambiguity of such a large percentage of graphemes introduces overlapping ambiguity: in 3,6% of cases, due to the fact that a grapheme can occur in all positions of a word (in the following example it is the second character, C2), a string consisting of three characters (C1C2C3) can be parsed both as C1C2 and C2C3. This is precisely because C2 can stand both at the beginning and the end of a word. Yen et al. (2012) tested whether the statistical probability of C2 representing a grapheme at the beginning vs. the end of a word when occurring in a target word C1C2 influences the reading process. They found that when C2 was a grapheme that is more frequently used as a grapheme at the beginning of words, participants had more difficulty parsing a target word C1C2 in which C2 occurred in the statistically less probable word-final position (cf. Yen et al. 2012: 1023).

Statistical considerations raise the question of which type of information readers of Chinese use in saccade programming, since saccade generation and specifically the factors that determine it in unspaced writing systems are still poorly understood phenomena (cf. Zhou et al 2018). While Yen et al. (2012) showed that statistical cues play a role in saccade programming, meaning that the probability of a grapheme to occur at either the beginning or the end of a word has an influence on how readers parse words, this effect could be verified mainly for foveal but not parafoveal processing.[254] The authors highlight that information both at the word and the grapheme levels affect eye movements: at the word level, word frequency influences fixation duration, as high-frequency words are fixated more briefly and are more likely to be skipped over than low-frequency words. Additionally, orthographic neighborhood size (which is measured in the number of words that are spelled similarly) and the size of the morphological family (= words sharing the same base word) affect fixations and saccade programming. By contrast, at the grapheme level, grapheme frequency and visual complexity of the graphemes' basic shapes influence eye movements (cf. Yan et al. 2006). Notably, now, information at the word level can be obtained only in the foveal region, while grapheme-level information can be obtained also in the parafoveal region. This means that in Yen et al.'s study cited above, an overlapping ambiguity (e.g. that C1C2C3 can be parsed as [C1C2]C3 or C1[C2C3]) that occurs in the parafoveal region does not interfere with subsequent reading performance (cf. Yen et al. 2012: 1027). This is consistent with findings from Thai, where no parafoveal-on-foveal effects could be found when the word frequency of the word to the right of the word that is currently fixated was manipulated (cf. Winskel & Perea 2014).

As for the question of where the eyes land for their fixations, investigations of a PVL in Chinese have led to mixed results and remain inconclusive (cf. Yen et al 2012: 1010). For example, Liu & Lu (2018: 346) claim that there exists no special landing position in Chinese, whereas Yan et al. (2010) propose a flexible eye guidance model: if sufficient information is obtained from the parafoveal region, the next saccade is targeted roughly in the middle of the word that will be fixated next, whereas if not enough information is obtained, the first grapheme, or

254 The perceptual span in Chinese covers three to four graphemes to the right of a fixation. This means that the foveal and parafoveal regions commonly span a two-grapheme foveal word, a two-grapheme parafoveal word, and one more grapheme (cf. Zhou et al. 2018: 730).

more generally, the beginning of the next word is fixated instead (but cf. Liu & Lu 2018 for a critique of this model).

As for Thai and Japanese above, it is interesting to ask what effect the introduction of word spacing has on users' reading processes in Chinese. Bai et al.'s (2008) main conclusion is that neither the insertion of word spaces nor the demarcation of word boundaries by other means (in their case greyhighlightingofeachotherword) facilitated reading or reduced reading time (but cf. Zang et al. 2013 who found that it does facilitate word recognition). A finding that Bai et al. deem equally important is that both modifications of the traditionally unspaced Chinese text – word spacing and grey highlighting – also did not influence the reading process negatively. This result, however, does not necessarily imply that the demarcation of word boundaries has no effect whatsoever in Chinese. It is rather likely that the facilitative effect provided by it is canceled out by the deleterious effect it has due to its unfamiliarity to Chinese readers (cf. Ma 2017: 820). This, again, highlights the conflict between the universal and system-specific levels of naturalness.

In a different study, Zhou et al. (2018) tested what effects another word boundary cue has on reading in Chinese: alternating color (which, granted, is similar to Bai et al.'s grey highlighting condition). When each other word is highlighted in color and this provides correct lexical segmentation cues, the landing position of saccades tends to be close to the center of words. By contrast, if the highlighting purposely provides an incorrect segmentation cue, the landing position shifts to the beginning of words. Interestingly, if no cue is given at all, which equals the default unspaced condition, the landing position is likewise closer to the beginning of words (cf. Ma 2017). In short, most studies focusing on word spacing in Chinese find that readers of Chinese can utilize parafoveal information about word boundaries for saccade generation in case such information is available.

A different crucial finding that Bai et al.'s (2008) study brought forth is that the word appears to be psychologically real for Chinese readers and that it is a more central unit than the grapheme or "character" (cf. also Hoosain 1992). Thus, it appears that empty spaces that demarcate words are not necessary for a unit to be of cognitive importance. Although the graphematic word does not exist as a structural unit in Chinese (at least not as defined for alphabets, cf. Section 2.5), it seems to be a relevant unit in processing.

An experiment in which L2 readers of Chinese were presented Chinese text with added word spaces (cf. Yao 2011) provides additional interesting findings concerning the interaction of the universal and system-specific levels of naturalness. In this experiment, readers of Chinese whose first writing system features word spacing (= readers of Arabic, English, French, Italian, Mongolian, Portuguese, Russian, Spanish, and Urdu) were not affected by the addition of spaces, whereas the reading process of advanced readers of Chinese whose first writing system lacks word spacing (= readers of Thai) was slowed down. This is striking insofar as the above-discussed studies on Thai suggest that the introduction of word spacing to Thai texts has no disruptive effect on the reading process for L1 Thai readers but instead even facilitated word processing (but not saccade programming). The fact that in Yao's (2011) study, Thai readers who were also advanced L2 readers of Chinese were negatively affected by the addition of spacing calls into question the universal naturalness of word spacing. According to the *Naturalness Differential Hypothesis* (cf. Schmid 1997: 337), in L2 acquisition, a feature of the L2 that is (at the universal level) more natural than the equivalent feature of the learner's L1 should be easy for L2 learners to acquire. However, Yao's (2011) findings contradict this. When native speakers (and readers) of Thai acquire Chinese as a second writing system, they do not have to grow accustomed to word spacing that is equally not a part of their first writing system; however, word spacing is claimed to be natural universally on the grounds of its benefits for eye movement guidance and word processing.

In their study, Winskel, Radach & Luksaneeyanawin (2009) presented Thai-English bilingual readers and English monolingual readers with unspaced English text. The authors' hypothesis was that the removal of spaces would affect Thai-English bilingual readers less since they are used to reading unspaced Thai. This would represent a positive cross-system transfer of reading strategies from Thai to English (cf. Genesee et al. 2006). Surprisingly, the results showed exactly the opposite: the bilinguals' reading was more severely disrupted than the reading of monolinguals. While the reading rate for Thai-English bilingual readers decreased by 45% when word spacing was removed, the reading rate for English monolinguals decreased only by 33%. This is in stark contrast to the findings of the study by Yao (2011) cited above in which the reading process of Thai readers with Chinese as a second writing system was disrupted by the introduction of word spacing.

A possible explanation for the apparently discrepant findings of these two studies is that when readers have reached a certain degree of proficiency in a writing system, the system-specific value of the feature [±spaced] is set, e.g. [−spaced] for Thai. Now, if the value of this feature is changed in an experimental condition, i.e. to [+spaced] for Thai and [−spaced] for English, this might disrupt reading regardless of the universally natural feature value, which is, arguably, [+spaced]. Of course, another possible explanation is that word spacing (or any form of demarcation of word boundaries, for that matter) is not actually natural at the universal level; however, given the evidence, this explanation is rather unlikely (see below). In sum, as established in Natural Morphology (cf. Section 4.2.1), system-specific naturalness can override universal naturalness, which both of the above-mentioned studies on reading in an L2 are in accordance with.

Another type of evidence pointing to the naturalness of word spacing for processing is provided by the history of writing. In short, the absence or presence of spacing in the history of writing reflects the diachronic development of reading processes. As Saenger (1991: 198), whose prolific work on the development of spacing culminated in Saenger (1997), noted about reading and writing: "it should not be assumed that these same cognitive activities have been used throughout human history". The central question, now, is why in some writing systems, during certain periods, word boundaries were not demarcated and text was written in so-called *scriptura continua* (i.e. unspaced text). This question is especially pressing given that a form of demarcation had existed prior to those periods.

As Saenger (1991: 207, emphasis in original) describes, in the earliest Greek inscriptions, words were separated by interpuncts, but Greece "soon thereafter became the first ancient civilization to employ *scriptura continua*". A lot later, the Romans followed, too. It is tempting to (rather drastically) discard this development – the elimination of useful information for word segmentation – as a "retrograde development in human history" (Saenger 1991: 208), although, under the perspective of the processing fit discussed here, this assessment could indeed be deemed accurate. However, factors that had nothing to do with processing were more dominant. They were of sociocultural nature:

> Stated summarily, the ancient world did not possess the desire, characteristic of modern civilizations, to make reading easier and swifter because the advantages, which the modern world perceives as accruing from easier reading – the swift effective retrieval of information in reference consultation, the ability to read swiftly a great many difficult technical, logical, and scientific texts, and the greater diffusion of literacy throughout all social strata of the population – were never or seldom viewed as advantages by ancients. (Saenger 1991: 208)

Saenger (1991: 209) further details that the "notion that the greater portion of the population should be autonomous [...] readers was entirely foreign to the elitist literate mentality of the ancient world", which is, arguably, socioculturally charged. Thus, as in many other cases, the sociocultural fit (cf. Chapter 8), specifically the social functions that writing fulfilled in these ancient societies, was more important than the processing fit, i.e. the legibility and readability of texts. Fittingly, the reintroduction of word spacing in the early Middle Ages by Irish and Anglo-Saxon scribes coincided with a shift in the relationship between readers and texts (cf. Saenger 1991: 210). Also, in terms of naturalness conflicts, it is paramount to mention that once word spacing had been reintroduced, the cognitive task of separating words was reassigned from readers to writers. That spacing, in the end, won out is in accordance with an assumed primacy of perception.

But how did the practice of reading text in scriptura continua affect reading practices and processes? Most importantly, reading at that time was not a silent but an oral activity. Reading aloud helped readers store in short-term memory the fragments of the text that had already been read, which allowed cognitive processes to be allocated to fragments not yet decoded (cf. Saenger 1991: 205). As Saenger (1991: 205f.) adds, readers were not entirely without help when reading scriptura continua given that scribes enlarged the spaces between individual graphemes. What is striking is the similarities between these ancient texts and texts designed for beginning (and often, oral) readers today. These modern texts, too, often place larger empty spaces between individual graphemes and additionally increase line spacing.

The fact that Greek and Latin are the best known instances of scriptura continua is no coincidence. Saenger (1991) describes how a fully (or almost fully) voweled segmental writing system – i.e. an alphabet or an abugida (cf. Section 2.7) – is a prerequisite for scriptura continua to function. By claiming this, he likely implies that a lack of both vowels *and* word spacing would render a writing system too complex for processing. Accordingly, ancient writing systems such as those of Mesopotamia, Phoenicia, and Israel "invariably contained separation" (Saenger 1991: 207). For Hebrew, Saenger (1991: 202) additionally mentions visual complexity as a factor that makes word spacing necessary, as even in the presence of both vowels and word spacing, the visual similarity of Hebrew basic shapes makes word recognition more difficult than in systems in which basic shapes are visually more distinctive.

## 7.3 Orthography

To very briefly illustrate how orthographic regulation can affect the processing fit, an example from Chinese shall be discussed. In Chinese, stroke order, i.e. the order in which the elementary forms of basic shapes are produced, is subject to orthographic regulation. This is odd from the perspective of most other writing systems, where the order in which strokes are written is completely free (from an orthographic perspective). Chinese stroke order rules are said to have originated in the production process and to have "been distilled from Chinese handwriting going back thousands of years", with the earliest evidence dating back as far as to 520 B.C. (Zhang 2014: 424). This situation is strikingly reminiscent of morphonological rules that still have a phonetic motivation described in in Natural Phonology (cf. Section 4.2.2): stroke order in Chinese, motivated by processing constraints, has likewise developed into a rule (or rather a set of rules). However, these rules still exhibit natural motivations grounded in production. Unlike morphonological rules, though, stroke order rules are not required for the production of well-formed units (in this case characters). Stroke order rules are thus located at a different level, the orthographic level.

What, now, is the use of such a regulation of the production process? Zhang (2014: 424, emphasis in original) points out that the prescribed stroke order "contributes to the correct, fast and aesthetic production of *hanzi*". This makes it unsurprising that it has become a central part of hanzi instruction. Furthermore, it is most commonly employed as an index key in Chinese dictionaries, meaning one needs to know the correct stroke order to be able to use dictionaries. In Section 7.1.2, natural processes of handwriting were identified in the form of the *grammar of action* (cf. Goodnow & Levine 1973). Is the orthographically prescribed stroke order sequence in Chinese similarly natural with respect to processing? Zhang (2014: 424f., emphasis in original) mentions evidence from neuropsychological studies that suggests it is:

> Research has demonstrated that enhanced performance was observed in different cognitive tasks when *hanzi* was presented as a sequence of strokes/radicals consistent with the standard writing sequence rather than in a random order [...]. The active role of stroke order in *hanzi* processing has been established as a motor schema, stored as part of the representation of *hanzi* in memory, and also as an effective aid in retrieving relevant information about *hanzi* from memory [...].

This example is presented here to show that orthographic regulation, while it does not introduce naturalness that was not there in the first place (in this case to the graphetic module), can at least enforce naturalness that already exists. This is the opposite of imposing unnaturalness that was not there. By regulating the stroke order that has been found to aid processing, orthography makes adhering to it binding for, as mentioned, dictionaries and instruction, i.e. textbooks, teaching material, etc. While due to its natural motivation, it is expected that the prescribed stroke order evolves naturally in children's writing acquisition, its status as an orthographic rule ensures that it does. For learners acquiring Chinese as a second language, stroke order is also relevant. While they are commonly unfamiliar with the concept of a regulated order of producing elementary forms, L2 users, too, adhere to implicit and internalized sequences of production. Studies show that the internalized sequences originating from their first writing system interfere with the orthographically regulated sequence in Chinese (in cases in which the two differ, that is), leading to production errors (cf. Zhang 2014). Thus, while there may be a sequence of production that is universally natural, there additionally exist sequences of production that have been established through users' experience with a given writing system (in most cases their first) and have, thus, become system-specifically natural.

# 8 Sociocultural fit

The processing fit focused on physiological and cognitive processing, which is an important aspect of how humans use writing. It is noteworthy that humans are usually oblivious to how they process writing – they just do. Consequently, how the systematic and linguistic fits interact with the processing fit is fairly unconscious. In other words, changes that are made to scripts and writing systems due to processing are not 'made' – they just happen, and usually this is a gradual process. By comparison, the sociocultural fit captures aspects of use that users are often a lot more conscious of. It is based on the fact that writing is a cultural technique that is used for communication in a myriad of ways that we can vaguely term 'scribal practices'. Crucially, it is not only *what* we write that communicates the messages we intend to convey. It is also the systems that we use to communicate that convey (additional) messages, whether we want them to or not. All the modules of writing systems – graphetics, graphematics, orthography – have an inherent indexical potential. They can signal, for example, parts of their users' identities by being connoted with a specific group of people (whether it is defined geographically, socially, …). This indexicality is something that users are often aware of, and it affects their acceptance of given writing systems. Against this background, it is unsurprising that historically, we find instances of people rejecting scripts, orthographies (especially after orthography reforms), or even entire writing systems. Since they are not only used as tools for communication but also for self-identification, writing systems are frequently perceived as common goods by members of literate communities. How the systematic and linguistic fits interact with the usage-based sociocultural fit, now, often depends on the explicit wishes of said communities. Consequently, changes that are made to writing systems on the basis of the sociocultural fit are predominantly based on conscious decisions. Importantly, they can override all other fits: to be accepted, a writing system does not necessarily need to suit the structures and features of a given language, and it does not even need to be processed efficiently; it does need a good sociocultural fit, however.

As in the previous chapters, dedicated subsections treat the distinct modules: Section 8.1 deals with graphetics, predominantly with questions of script choice (and switches) and biscriptality, while Section 8.2 illustrates how the demands and needs of literate communities shape the graphematic makeup of writing systems (among other things); finally, Section 8.3 turns to orthography, the module for which sociocultural considerations are arguably most important. Accordingly, it investigates sociocultural aspects of the establishment and reform of orthographies.

## 8.1 Graphetics

Writing is a cultural technique, so it is unsurprising that sociocultural considerations ought to play a central role in grapholinguistics. Since it is perceivable directly with the visual and/or tactile sense(s), the material appearance of writing is of particular importance to members of literate cultures – especially since it can potentially convey something about them. In this vein, Klinkenberg & Polis (2018: 82) note in their semiotic approach that the appearance of writing "is frequently associated with a producer or a group of producers [and] yields a signature effect". Corresponding to the different types of signs, it can assume symbolic, indexical, or iconic functions. This applies to the appearance of a given person's handwriting, but also to someone's choice of typeface in a given context. At a more global level, it also applies to the appearance of an entire script that is used to materialize the writing system in a given literate culture. Writing represents not only language – it represents much more.

'Sociocultural' is used here as an umbrella term. While it certainly does not collect all remaining aspects that cannot be reasonably categorized as belonging to the systematic, linguistic, or processing fits, it does bundle the many aspects that can be broadly classified as being of social and cultural relevance. This includes, among others, political, religious, and technological aspects. Notably, the sociocultural fit of scripts is inseparably tied to the sociocultural fit of writing systems that employ these very scripts, much more so than the systematic and processing fits of scripts are associated with the linguistic and processing fits of their corresponding writing systems. Thus, while for these latter fits, the distinction between the materiality of writing and the linguistic functionality of writing warrants individual treatments, with respect to the sociocultural fit, the material cannot be clearly separated from the linguistic. One of the reasons for this is that sociocultural factors, not unlike processing factors, are best analyzed from the perspective of users – everyday users, that is. These users, unlike linguists (though, admittedly, not even all linguists), usually do not distinguish between scripts and writing systems. Lacking the necessary meta-grapholinguistic knowledge, users are largely oblivious to the commonalities and differences between various types of writing systems and their respective linguistic functions – instead, what they do perceive, i.e. *see* as writing, more specifically as *their* writing, is the script they use.[255] Consequently, the connection between the graphetic and graphematic sociocultural fits is so strong that it is not always possible to straightforwardly determine whether a given factor concerns graphetics, graphematics, orthography, or all three of them. In this section, the main focus will be on those sociocultural aspects that are influenced mainly by the materiality of writing, whereas broader sociocultural questions that concern both the graphetic and graphematic modules will be discussed in the next section.

255 This is why many people would probably say that English, German, Spanish, Finnish etc. use the same "writing system" – because the writing systems of these languages all use the same script.

The arguably most pressing sociocultural question with respect to graphetics arises in the creation of writing systems for yet unwritten languages. Specifically, it concerns the major decision of how the new writing system should look – and *why* it should look that way. At this stage, the central question is whether an existing script of a different writing system is adopted or an entirely new script is devised. This also subsumes the question of which materials and technologies are even available for writing in a given (geographical, political, …) context, which can have major repercussions for the material makeup of scripts. To follow a chronological order, in this section, I will first address sociocultural issues that pertain to the creation (and, partially, reform, which represents a 'recreation' of sorts) of writing systems, after which I will turn to the issues that come to the fore once an established writing system is in use.

Unseth (2005) claims that the visual appearance of a new writing system is always of critical importance, whether a non-literate community for which the new writing system is developed wishes to distance itself from a (most often more dominant) community or, by contrast, wants to convey membership to a different literate community (these strategies are more closely analyzed in the next section). Accordingly, two of the most common wishes that communities express in the process of literacy development are "We want the orthography to 'look like' another language" or, inversely, "We do not want the orthography to 'look like' another language" (Cahill 2014: 13f.).

Very often, the first decision that is made is to adopt Roman script. This decision is frequently influenced by hegemonic and technological factors (see below). Consequently, this choice of Roman script can affect, bottom-up, the graphematics of the newly devised writing system. For example, it appears reasonable that the type of writing system most closely associated with Roman script be adopted along with it: the alphabet. And most often, it is not only the type of writing system that is adopted but also many of the individual graphematic values that basic shapes have in a given donor writing system (e.g. English or French). At this point, it is imperative to note that although the adoption of a script and the adoption of graphematic relations (i.e. relations between shapes and linguistic units such as phonemes) are often associated, these processes are in principle completely independent of one another. Thus, a script such as Roman script can potentially be adopted and utilized for a different type of writing system than the one it is mainly used for. Take, for example, the Roman basic shapes used in the syllabographic writing system of Cherokee (cf. also the introduction to Part II).

When instead, the decision is made to devise a new script from scratch, the above-mentioned wishes of signaling distance or affiliation also come to the fore, meaning that a new script can also be greatly influenced by (an) existing script(s). An apparent graphetic example is the "stroke formation" of basic shapes "in the Sinitic sphere" (Unseth 2005: 33), mainly Chinese, Japanese kana, and Korean Hangul. The basic shapes in these scripts share visually salient features, and outsiders not literate in any of the writing systems in which these scripts are used could be led to believe that they belong to the same set (cf. Figure 13). This can be at least partially attributed to the fact that script creations from “scratch”, such as the creation of Korean Hangul, do not occur in a geographical, sociocultural, and political vacuum. Instead, linguistic communities make decisions about the design of a new script “with a conscious awareness of their neighbors” (Unseth 2005: 33). In the cited case of Japanese, for example, the visual appearance of the basic shapes in the kana inventories is not arbitrary but a result of the fact that they developed diachronically out of Chinese basic shapes (cf. Takagi 2019; Smith 1996: 210–213).

Crucially, the choice of script is socioculturally charged because it is intimately tied to the question of how a community sees and identifies itself. As Abdelhay, Makoni & Makoni (2018: 98) observe,

> script choice is not a neutral corpus-planning endeavour but, rather, is deeply associated with theological and political issues to the extent that script choices are often oriented towards the production of binary, socially fragmented spaces of identification.

The history of writing Azeri serves as an example of a complex trajectory of script changes that reflects how scripts can be instrumentalized as (political) tools for identification. In fact, the script for the writing system of Azeri was changed three times in the 20th century alone. The first important script change was a switch from “twelve or more” (Bahadori 1993: 10) ancient Azeri scripts to Arabic script, which occurred in the 7th century when Arabs arrived in Azerbaijan. This first script change was driven by cultural and religious motivations and had the effect of “reinforcing Azerbaijan’s link with Islam” (Hatcher 2008: 107). Crucially, “Arabic script presented many difficulties, not effectively representing Azeri phonetically” (Hatcher 2008: 107), which highlights how the linguistic and sociocultural fits can conflict. It also underlines that the switch to Arabic script was actually not a pure script adaptation, but, as illustrated above, an adaptation of the type of writing system – an abjad –, and along with it, Arabic graphematic relations. The first major script change of the 20th century came in 1924 when, due to a new Soviet policy, Arabic script was dropped and replaced by Roman script. Although this decision was promoted as part of a liberal language policy based on the premise that “everyone had the right to speak whichever language they wanted, privately or publicly” (Hatcher 2008: 107) and was supported by the intelligentsia of Azerbaijan, it had an ulterior political motive: to distance the people of Azerbaijan from Islam in order to “secularize Azerbaijani identity” (Hatcher 2008: 109). Quickly, the supposedly liberal character of the language policy faded, and in 1925, a decree “outlawed the importing of anything printed in Arabic script” (Hatcher 2008: 108). By 1928, in many villages, Arabic books were being burned, and owning or hiding them could result in imprisonment or worse (cf. Hatcher 2008: 109). Then, in the 1930s, the once progressive language policy of the Soviet regime changed even further. A crucial, politically motivated event that took place in 1928 and exerted an influence on the situation in Azerbaijan was the Turkish writing system switching from Arabic to Roman script (cf. Wood 1929). Following this change, the Soviet regime started to fear a pan-Turkic identity movement, and as a result, in 1939, Stalin announced that for the Turkic languages of the Soviet Union, Cyrillic script would replace Roman script. This marks the second major 20th-century script change for Azeri. The main goals behind it were “Russification and isolation between Turkic nations” (Bahadori 1993: 11). Finally, the third script change occurred shortly after Azerbaijan gained independence in 1991: only four days after signing the Alma Ata Protocol that dissolved the Soviet Union, the parliament in Azerbaijan voted for a return to Roman script (cf. Hatcher 2008: 111).

As this turbulent history of script changes shows, scripts, and writing systems in general, are tools that are instrumentalized by political forces. This is precisely because scripts carry in them deep cultural connotations. They are not just inventories of shapes – they always have their own history and are connected to cultures, religions, politics, nations, and much more.

The sociocultural fit is affected also by technology. Accordingly, the question of which technology (materials, instruments, etc.) is available to a given community is by no means trivial. On the contrary, like other sociocultural factors, it can prove to be a decisive factor in the material makeup of a script. Burmese script is a telling example (cf. Section 1.1) that highlights the relevance of the paragraphetic perspective. At the time when the script was developed, what was available as a writing surface, and what was ultimately *chosen* to write on, was palm leaves. The round visual appearance of the script's shapes is a consequence of this choice. Even though the preliminary conclusions discussed in the context of the processing fit of scripts suggest that curves might be easier to process than angles and straight lines (cf. Section 7.1.1), cognitive and physiological human pressure was not the driving force behind the round shapes of Burmese – the material was. Availability of suitable writing material is driven by geographical and supply factors, as what is relevant is what material occurs (naturally) in a given region. After all, in the design of new scripts, script creators can only work with what is available. Similarly, when existing scripts are adopted, scripts for which sufficient technology is available or for which existing technology is suited are more likely to be adopted.

Not only technology in the sense of available materials is of central concern for the sociocultural fit but also the degree of a community's technological advancement, and in that context, the availability of modern technology such as computers, keyboards, fonts, etc. These technologies shape modern literacy development to a tremendous degree, which is reflected, for example, in publications detailing the efforts of the *Summer Institute of Linguistics* (SIL) to bring literacy to non-literate communities. Accordingly, among the larger group of “non-linguistic factors”, Cahill (2014: 9) highlights technical issues which “include printability and Unicode-compliance” of basic shapes. One of the likely reasons why Roman script is so popular in literacy development is that communities who gain literacy through a new writing system that employ it can immediately make use of all of the technology that is suitable for Roman script – as long as the instruments are available. This is not reduced to physical instruments such as keyboards and smartphones but includes also the internet and social media, where, even though many other scripts can be used, Roman script is frequently employed, even for writing systems that are written with other scripts (think of Romanizations of Greek, Arabic, Thai, etc.). In the end, for a writing system not only to be accepted by a community but also to come and actually remain in continuous use, "[t]here needs to be a stable mode of production to enable local people to continue with literacy on their own" (Cahill 2014: 21).

In this context, one could argue that the option of creating new scripts from scratch is nowadays often not even considered given that new shapes – even if they exhibit natural systematic and processing fits – are not immediately suitable for technology as they are not yet encoded in Unicode. This cuts off many if not all possibilities of using modern digital forms of communication. This appears counterproductive when one considers that the main goal of creating a writing system is to enable written – and nowadays this includes digital – communication.

A situation that reveals a whole bundle of factors that are relevant for the sociocultural fit is a group's (or individual's, for that matter) choice between two or more coexisting scripts, visual varieties of scripts, or orthographies. Bunčić, Lippert & Rabus (2016) subsume these socioculturally charged choices under the heading of *biscriptality* and define it as "the simultaneous use of two (or more) writing systems (including different orthographies) for (varieties of) the same language" (Bunčić 2016a: 54). In an effort to clarify, organize, and operationalize concepts and terminology central to the sociolinguistics of writing, Bunčić (2016a) draws up a typology in which various types of biscriptality are characterized (cf. Table 10).

Based on the type of opposition – in the Trubetzkoyan sense – between two scripts, Bunčić assumes privative and equipollent situations. In (1) *digraphia*, there is a privative opposition between scripts, meaning one script is lacking a feature that is exhibited by the other script. Which of the two scripts is used in given situations is determined by (1a) *diaphasic* (pertaining to registers and style), (1b) *diastratic* (pertaining to social strata), (1c) *diamesic* (pertaining to the conceptual dimension of written vs. spoken established by Koch & Oesterreicher 1985), or (1d) *medial* (depending on the writing material) factors.

By contrast, an equipollent opposition is present in cases in which two scripts are characterized by the presence of two distinct features, respectively, i.e. the feature [Hindi] for Devanāgarī as used in the Hindi writing system and [Muslim] for Arabic script as used for the writing system of Urdu (cf. Bunčić 2016a: 55). This type of biscriptality is termed (2) *scriptal pluricentricity*. It is determined by (2a) *diatopic* (i.e. geographical), (2b) *ethnic*, or (2c) *confessional*, i.e. *religious* factors.

Finally, there are also more complex situations in which no clear-cut criterion can predict the choice of script. Bunčić (2016a: 60) calls these situations *diasituative* and names Serbian as an example, where a variety of factors (e.g. "the number of participants in a communicative setting; the relationships among participants concerning age, education, sex, etc.; time and duration of the communicative act; the topic; the degree of publicity", Bunčić, 2016a: 61) can influence whether a writer will choose either Roman or Cyrillic script. This type of biscriptality is referred to as (3) *bigraphism*. It is crucial to note that "scripts used in a bigraphic situation always carry additional indexical meanings" (Bunčić 2016a: 61).

This typology of biscriptality is based on sociolinguistic concepts. In what is a truly fine-grained model, Bunčić also proposes a grapholinguistic dimension that differentiates between the graphetic, graphematic, and orthographic levels.

All of the types of biscriptality given above concern scripts, as digraphia, scriptal pluricentricity, and bigraphism always involve two or more distinct scripts such as Devanāgarī and Arabic script. Scripts are treated by graphetics, which implies that these types of biscriptality are located at the graphetic level. However, from the perspective of users, digraphia, scriptal pluricentricity, and bigraphism are, in fact, graphematic: Bunčić (2016a: 64) illustrates that in a digraphic situation, for example, users might not be able to read both of the scripts that are in a private opposition. The substitution of one script for another is not merely a change of basic shapes but instead a fundamental alteration of both the visual appearance and linguistic makeup of the entire system which renders it "indecipherable" to readers who do not read the 'other' script. In this context, Bunčić (2016a: 64) argues that "different scripts completely block communication". Take as an example the script changes in Azerbaijan discussed above: switching the script from Arabic to Roman not only altered the graphematic relations between basic shapes and phonemes but, in fact, the type of writing system: an abjad became an alphabet (see above for the argument that scripts are usually adopted together with their common graphematic features). Thus, in such a situation, users of a writing system do not have to just familiarize themselves with new basic shapes that have the same graphematic relations as the old ones. Instead, they must learn a new system altogether. Such situations in which not only a script but, with it, the type of writing system is changed, are referred to as *intersystemic biscriptality*. However, it is important to note here that in the case of Azeri, no two scripts were in use simultaneously, at least not officially and for a long time, as one script was always replaced by another by decree. Thus, the case of writing Azeri is actually no fitting example of biscriptality, which designates a situation in which scripts always coexist simultaneously. Nonetheless, it highlights that changing or switching scripts is not a graphetic matter but a graphematic one. Notably, this also holds for a substitution of scripts that does not involve a change in the type of writing system, such as when Cyrillic is replaced by Roman, which are both used alphabetically. This situation is referred to as *intrasystemic biscriptality*.

Bunčić (2016a: 64) does also propose a graphetic level of biscriptality, which he calls the glyphic level. Its subtypes are, parallel to the types characterized above, *diglyphia*, *glyphic pluricentricity*, and *biglyphism* (cf. Table 10). In Section 1.2.1, *glyph* was defined as a synonym of *graph*, the concrete materialization of a basic shape. Graphs (or better: graph classes, cf. Section 2.3) can exploit the graphetic solution space of a given basic shape and are the locus of variation that gives rise to different styles of handwriting or typefaces. *Blackletter* and *roman*, for example, are glyphic variants of Roman script. Unlike the distinct scripts in the situation described above, as glyphic variants of one script, they can be read by users familiar with said script. Glyphic variants that are of sociolinguistic relevance are typically not just different typefaces (such as Arial vs. Times New Roman) but variants that are perceived as (almost) emblematic of a given sociolinguistic variable; examples follow below.

The third grapholinguistic level of biscriptality is the orthographic level. It distinguishes different types of standardization: an example is the variant <ss> used in the German variety of Switzerland and Liechtenstein as opposed to <ß> which is used in other German-language varieties. The subtypes of the orthographic level are, unsurprisingly, *diorthographia*, *orthographic pluricentricity*, and *biorthographia*.

Table 10 provides an overview of the sociolinguistic and grapholinguistic dimensions of biscriptality and offers examples taken from Bunčić, Lippert & Rabus (2016).

Table 10. Types of biscriptality, adapted from Bunčić (2016a: 67) with examples from Bunčić, Lippert & Rabus (2016)

| **grapholinguistic dimension** / *sociolinguistic dimension* | **glyphic variant** GRAPHETIC | **script** GRAPHEMATIC | **orthography** ORTHOGRAPHIC |
|---|---|---|---|
| *privative* (diaphasic, diastratic, diamesic, medial) | **diglyphia** diaphasic diglyphia in Russia | **digraphia** diamesic (medial) digraphia in medieval Scandinavia | **diorthographia** medial diorthographia in Novgorod |
| *equipollent* (diatopic, ethnic, confessional) | **glyphic pluricentricity** Orthodox, Muslim, and Catholic Cyrillic in Bosnia | **scriptal pluricentricity** Hindi vs. Urdu | **orthographic pluricentricity** Simplified vs. traditional Chinese |
| *diasituative* | **biglyphism** Blackletter and roman in German | **bigraphism** Cyrillic and Roman for Serbian and "Serbo-Croatian" | **biorthographism** 'Classical' and 'Mistralian' spelling in Occitan |

In the discussion of the creation of new writing systems, it became clear that the choice of a script is complex and subject to a variety of factors, many of which are sociocultural. However, in these cases, usually, once a choice is made, it is expected to be absolute, and no situations of biscriptality emerge. However, the complex sociolinguistic situations of biscriptality as described in Bunčić, Lippert & Rabus (2016) are of special interest since the analysis of the choices users make offers evidence for categories relevant for evaluation and specifically the degree of sociocultural fit. In situations of biscriptality, distinct variants occur simultaneously in *one* language. Despite their *simultaneous* occurrence and the prevalent use of different variants, it would be wrong to assume that users of these writing systems (always) have a free choice. As is evident from the table above, the choice of glyphic variant, script, or orthography is influenced by a variety of factors. Diasituative situations are exceptions as they are characterized by the possibility of a more or less free choice, even if in reality, they are also constrained by a bundle of relevant factors. It is the difficulty of identifying which of them is the deciding factor that can lead to the assessment that, for example, the choice of a Serb to use Cyrillic script is "free" when in fact it is not. In sum, all of the factors listed in the sociolinguistic dimension column in Table 10 are central in the investigation of the sociocultural fit of scripts.

As an example, consider *medial biscriptality*. Here, it seems obvious which factor affects the choice: choosing a given glyphic variant, script, or orthography is determined by the respective material used for writing. Take medial diorthographia in Novgorod, where birch-bark documents were written in a vernacular orthography while religious books used a standard orthography. The digraphia involving Latin script and runes in medieval Scandinavia might also appear to be determined by the writing material: while Latin script is written with ink on parchment, runes are carved on wooden sticks or bones. However, this case is actually more complex, as the choice of material is itself determined by a different factor: the intended function of the text. For texts that were meant to last, scribes chose parchment and ink, which simultaneously determined that they use Latin script. By comparison, for rather ephemeral texts that could be discarded after reading, wood or bones were used – and, consequently, runes. Thus, a fine-grained analysis of the contexts in which these two scripts were used reveals that Latin script exhibits the feature [+distance] that runes lack as they are associated with "spontaneity, intimacy with the addressee, expressiveness or privateness" (Bunčić 2016b: 76). In a nutshell: the digraphia between Latin and runes is not actually medial but diamesic.

Although it would be useful for a theory of writing, an identification of specific categories for the sociocultural fit of scripts as was made for the other fits is a challenging endeavor. The sociolinguistic dimensions listed in Table 10 appear to be promising candidates but are not themselves categories. Instead, they are factors that add up to and affect a more general category affecting the degree of a script's sociocultural fit – one that could be called *situational adequacy*. 'Situational' implies that a script's sociocultural fit is not absolute but must rather be evaluated individually and anew for each given situation. This is where the multitude of factors – such as religious or ethnic factors – outlined above comes into play, all of which affect what is perceived as situationally adequate. What is also important to note is that while these factors are finite and similar for each situation, their prioritization differs for each given situation. In one situation, technological factors such as the materials available might be prioritized and render an angular script carved in wood most situationally adequate, while in a different situation, even if wood was the only material available to serve as a writing surface, adopting Roman script could still be considered situationally most adequate due to other –

e.g. political – factors that override technological factors. This all points to the fact that the sociocultural fit of a script must always be evaluated at the system-specific level – and even there it might be dynamic depending on the given situation.

For the sake of completeness, I want to mention briefly that the graphetic sociocultural fit cannot only be evaluated with respect to scripts as inventories of basic shapes but also with respect to visual variation at the level of graphs, which corresponds with Bunčić's (2016a) inclusion of the level of glyphs. An obvious question at this level is whether a specific typeface is situationally adequate, such as Comic Sans in PowerPoint presentations of university professors (cf. Meletis 2020a). Not the systematic or processing fits of Roman script are central to this question but, taking into consideration the degree of possible visual variation within the graphetic solution spaces of the script's basic shapes, the specific appearance of the basic shapes of Comic Sans – complete with the connotations that it evokes. These connotations restrict the contexts in which using Comic Sans is deemed either adequate or inadequate. Questions like these add a layer of complexity since the situational adequacy of a typeface cannot be evaluated by a descriptive grapholinguistic analysis alone – nothing in the appearance of Comic Sans makes it inherently adequate or inadequate for given situations. Instead, (implicit) conventions, norms, and attitudes that users in a literate community negotiate among themselves come to the forefront, which is why this question will be addressed more elaborately in the context of the sociocultural fit of orthographies in Section 8.3.

## 8.2 Graphematics

> […] literacy must respond explicitly to the needs of the specific speech community involved. (Jones & Mooney 2017: 5)

While the sociocultural fit of scripts (see above) is concerned with how the visual appearance of a writing system can suit sociocultural needs (or not), the sociocultural fit of the graphematic module subsumes the sociocultural fit of scripts but simultaneously opens up a much larger sociolinguistic perspective. In a nutshell, it evaluates whether an entire writing system suits a given literate culture. To answer this question, a culture's complex context needs to be taken into consideration, including historical, geographical, etc. factors. The sociocultural fit is determined strongly by external needs which result from the status that writing, as a cultural technique, occupies in society in general and, at a more local level, the (varying) status it occupies in specific literate cultures.

Writing systems are frequently instrumentalized as tools to fulfill political functions (cf. Coulmas 2000: 47). For this reason alone, it would be fatal to restrict a theory of writing to the linguistic and processing fits. As Weth & Juffermans (2018: 7, my emphasis) put it: "In interaction with humans, writing does a lot more than represent speech [or rather language, D.M.] visually. The technological agency of writing turns out to be not a transcription system and memory record but *a tool for solving social problems*". As such a tool, writing represents "a symbolic resource with unique affordances for the construction of languages and projects of self-identification" (Abdelhay, Makoni & Makoni 2018: 97). In this section, I will outline some of the main considerations that factor into the sociocultural fit of the graphematic module and writing systems in general. As in the treatment of the sociocultural fit of graphetics, what will take center stage here is the creation of writing systems for yet unwritten languages, a process that proves illuminating with respect to sociocultural factors and has conveniently spawned an abundance of literature elucidating literacy development from various perspectives. Issues that will also be addressed are the possibilities of sociocultural variation within already established writing systems and the interaction of said variation with the linguistic and processing fits.

This is not the place to discuss the general relationship between *literacy* vs. *orality* and literate vs. oral cultures (for the seminal study on this topic cf. Ong 2012; for an overview cf. Dürscheid 2016: 53–61). However, it is paramount to keep in mind that literacy and orality are often not treated as equal states but as endpoints of a development, with orality as the original state and literacy as the more advanced goal. Thus, orality and literacy are hierarchized, and in the context of this hierarchization, "writing can be critiqued [...] on ideological grounds as a tool used by human beings" (Weth & Juffermans 2018: 1). Statements like this invite a critical inquiry of *how* writing is used as an ideological tool. In this vein, aside from evaluating different writing systems and comparing them with respect to how well they are suited for the needs of the literate cultures they are used in (which is done in the following), one could also compare the ideological undertones of orality vs. literacy. In their introduction to a volume titled *The tyranny of writing* (a title referring to a statement Ferdinand de Saussure made about writing), Weth & Juffermans (2018: 3) note that "literacy is still very unequally distributed in the world". Illiteracy, which was "once a normal state of affairs[,] is now considered a disadvantage and an obstacle to human and social development" (Coulmas 2003: 225). These statements, however, refer mainly to the differing levels of literacy in literate cultures and claim that to be illiterate in these cultures represents a disadvantage. The more global difference between literate and oral cultures is, however, arguably also used for similar evaluative classifications. It is telling that oral cultures are often seen through a lens of deficiency and referred to as illiterate cultures, reflecting that the perspective taken is that of the chronologically secondary phenomenon – literacy, which has to develop out of orality, however (cf. Dürscheid 2016: 55). Accordingly, whether a language has or lacks a writing system is a factor that is believed to hierarchically distinguish it from other languages. Even within a single language, the variety or 'dialect' that is chosen as the basis for the written standard is perceived as being of a preferred status compared to other varieties of the language (see below).

Writing is all about social choices. These choices can be neutral or socially charged (cf. Sebba 2009: 38). Frequently, they are indeed socially charged when, as mentioned above, writing is used as an ideological tool as "human beings […] individually and collectively use or abuse writing in making social distinctions, intentionally and consciously or otherwise" (Weth & Juffermans 2018: 7). One of the fundamental features of writing that allow such an instrumentalization is its inherent normativity. In this respect, the effect that writing has on the spoken language in literate cultures must not be underestimated: the evaluation of correct or 'wrong' pronunciations, or, at the syntactic level, the evaluation of good spoken "sentences" (cf. Dürscheid 2016: 57) demonstrably reflect the transfer of literate norms to the spoken language (cf. also Karg 2015: 28–33). The hybrid functional nature of both writing and speech is captured by a conceptual distinction that has been impactful in the German-speaking realm: Koch & Oesterreicher's (1985, 1994; for an English translation, cf. Koch & Oesterreicher 2012) continuum of *orality* and *literacy* (cf. also Biber 1988). In their conception, the dimension of *medium* – whether a text is medially, i.e. materially, realized in the spoken or written modality – is divorced from the *conceptual* dimension. This distinction allows acknowledging that, for example, medially spoken texts can be conceptually written such as academic presentations at conferences: medially, they are spoken, but the means of expression used in them – their structural makeup, register, and style – are characteristic of written language and thus mark them as conceptually written. The inverse is true for many types of informal CMC (computer-mediated communication): consider text messages in messenger services such as WhatsApp. They are often conceptually spoken although they are commonly medially written (for a more thorough discussion of Koch & Oesterreicher's concept, cf. Dürscheid 2016: 43–53). In any case, the dynamic interaction between the medial and conceptual dimensions proves that there is no 'great divide' between literacy and orality. Instead, there is a continuum.

The impact of *autonomous literacy*, i.e. reading and writing as technical skills divorced from any social context, must not be overestimated: What Goody & Watt (1963) proclaimed as consequences or, less neutrally phrased, *benefits* of literacy – the ability to think analytically, abstractly, and logically, to specialize and compartmentalize knowledge, to store and chronologically order knowledge, among others – might not be direct consequences of literacy *per se*. In fact, "the view that, in order to become socially and cognitively equal, preliterate societies must become literate has […] been discredited" (Jones & Mooney 2017: 4). A "concept of 'traditional literacy' […] associated […] with 'formal, Western-style' education" is thus "too restrictive and inapplicable" for the socioculturally diverse contexts of literacy development (Jones & Mooney 2017: 4). This is in line with what Scribner & Cole (1981) suggest on the basis of their study in which they focus on non-formal education: the above-mentioned 'benefits' are not directly related to literacy as a technical skill but are instead consequences of education. This allows the conclusion that "context-independent abstract thought, memorization skills and logical thinking are actually more correlated to schooling and urbanism than literacy as the mere ability to read and write" (Weth & Juffermans 2018: 9). What matters, thus, is not (just) the acquisition of literacy but the social circumstances in which it is acquired and that in any case, literacy is not interpreted only as a technical skill but "involves the acquisition of knowledge about how to apply these skills in specific contexts and for specific purposes" (Jones & Mooney 2017: 4). These conclusions are emphasized by the paradigm of *New Literacy Studies* (cf. Street 1984; Barton 2007) which focuses on the sociocultural aspects and contexts of literacy.

Another approach closely linked to Scribner & Cole's investigation of non-formal education is the study of so-called *grassroots literacy*, which is defined as "writing produced under poor material and infrastructural conditions and at a distance from the institutions of prescriptive and elite-linguistic normativity" (Weth & Juffermans 2018: 10). Not unlike writing systems in the past, at a time when no prescriptive codified orthographies had been put in place (cf. Chapter 3), grassroots literacy gives rise to a high level of graphematic variation due to the lack of orthographic regulation. Cultures with this type of non-prescriptive literacy are thus free of many of the socially charged tensions that persist in more elaborated literate cultures in which normativity has been established. This underlines that literacy is not an inherently ideological tool but that it is instead instrumentalized as such by people. Orthography and the prescriptivism associated with it are at the core of this instrumentalization. Despite counterexamples such as grassroots literacy, literacy is, in general, not divorced from social 'baggage'. Thus, it is not defined "as some decontextualized 'ability' to write or read, but the social practices into which people are apprenticed as part of a social group, whether as students in school, letter writers in a local community, or members of a religious group" (Gee 2008: 80).

To conclude this introductory excursus, it suffices to say that attitudes towards literacy always appear ideologically charged. This becomes obvious when literate societies are equated with 'modern' societies or when illiteracy is, as mentioned above, interpreted as an indication of insufficient education, failure, or economic disadvantage (cf. Coulmas 2003: 223). These considerations will not play a prominent role in the following, in which the focus lies on operationalizing the comparison of literate cultures and their writing systems. However, since the sociocultural fit of writing systems is most crucial in the process of creating new writing systems, also known as literacy development, the mentioned attitudes associated with literacy and orality (or 'illiteracy') should be kept in mind.

Previous lists of criteria compiled to evaluate writing systems (cf. Section 4.1) almost always include social, cultural, political, etc. considerations of some kind. Often, these are collected under the label of "sociolinguistic factors", where 'sociolinguistic' "can be taken as a broad term covering a range of social, cultural, and historical aspects" (Sebba 2009: 36). Together with psychological and pedagogical considerations, they are sometimes subsumed under the heading of "non-linguistic factors" (cf. Seifart 2006; Cahill & Karan 2008; Cahill 2014), corresponding with the present distinction between the systematic and linguistic fits on the one hand and the processing and sociocultural fits on the other. These sociolinguistic factors include top-down factors such as governmental policies, which, however, more often regulate orthographies than writing systems (for the difference in regulation, see below). They also include factors such as dialects, relationships with and attitudes towards other literate cultures, the transferability of a given writing system to other languages, and technological factors such as the material available and the question whether a script used for a writing system is encoded in Unicode (cf. also the graphetic sociocultural fit above).

At the outset of the discussion of these sociocultural factors, it must be noted that they cannot be measured as straightforwardly as the categories relevant to the linguistic or processing fits. Sociocultural factors are often very subjective in nature and must therefore be assessed individually. Furthermore, what has a high sociocultural fit for a given culture can often not be decided from the outside – frequently, it is rather a matter of internal negotiation. Accordingly, grapholinguists investigating these aspects must imperatively consider the discussions that are internal to a given culture. Despite its fuzzy and idiosyncratic-seeming nature, the sociocultural fit is arguably the most important of the four fits. In fact, the success or failure of a writing system hinges on it, as Cahill (2014: 16, emphasis in original) underlines:

> People accept or reject an orthography based on sociolinguistic factors. If a group doesn't *want* to use an orthography, it doesn't matter how linguistically sound it is – they won't use it. So 'what the people want' is not just one more factor; it is the *most critical* factor in *acceptance* of an orthography.

In other words, a writing system that is not accepted by its community has effectively failed (cf. also Jones & Mooney 2017: 23; Hinton 2014: 144). Crucially, what the members of a literate culture who commonly lack meta-knowledge about writing systems usually judge is precisely the palpable sociocultural fit of a writing system and not its linguistic or processing fits. Since the sociocultural fit is so specific to given literate cultures, general categories may be identified for it but it is impossible to proclaim what has, universally, the best sociocultural fit precisely because what is socioculturally suited always depends on the needs and wishes of a specific community. Accordingly, the sociocultural fit, in general, is located at the universal level, but the specific sociocultural fit of given writing systems is located at the system-specific level.

As mentioned above, today, literacy is widespread, but it is still far from being ubiquitous. There exist many communities whose languages have no writing system. This is the context of so-called *orthography development* (cf. Lüpke 2011), also referred to as *graphization* (cf. Jones & Mooney 2017: 1) or – the term preferred here – *literacy development*. The creation of new writing systems for unwritten languages is central to an investigation of the sociocultural fit as it reveals the sociocultural factors that are discussed in this section 'in action'. Indeed, in many creations of writing systems described in the literature, sociocultural factors prove pivotal. When compared to *ancient grammatogenies*, i.e. the few instances in which writing was created from scratch thousands of years ago, more recent creations of writing systems in the context of literacy development are of special interest since they occur in "controlled" environments. Usually, linguists or scholars with other areas of expertise are on scene and work together with the language community in question (or sometimes, they work without involvement from the community) on the task of providing a language with a writing system. In most cases, thus, there is an outside specialist of some kind who collaborates with community members, i.e. people on the 'inside'. Outsiders involved in literacy development are referred to as *script mediators*, *orthography mediators*, or *orthographers* (cf. Sebba 2009: 41; Jones & Mooney 2017: 1); many of them are or have been members of the *Summer Institute of Linguistics*. They have the chance to describe sociocultural factors as they encounter them during their work, rendering their accounts of the influence of sociocultural factors on a writing system's makeup not external *a posteriori* explanations but invaluable first-hand descriptions. This is also the reason I will focus on these instances of literacy development in most of the remainder of this section.[256] Conversely, for older grammatogenies in which no outsiders were involved (which, however, are far fewer, cf. Sebba 2009: 41), often only attempts can be made at historical reconstructions of circumstances that led to specific choices.

Following Fasold's (1984) description of factors relevant in the choice of (national) languages, Unseth (2005) lists three sociocultural factors that play a crucial role in the creation of a writing system: (1) identification with a group, (2) distancing from a group, and (3) participation in developments on a larger scale. Smalley (1964) proposes different factors. As mentioned in Section 4.1, his criteria are meant as evaluation devices for assessing the quality of a writing system and, consequently, comparing different writing systems. Of his five criteria, three are of sociocultural nature: (I) maximum motivation for the learner and acceptance of a writing system by society and controlling groups such as the government, (II) maximum transfer, and (III) maximum ease of reproduction. While Unseth's categories give this section its overall structure, Smalley's categories will be interspersed at relevant points.

One of the central factors in the creation of writing systems, and the first one listed by Smalley, is (I) maximum motivation for the learner. Arguably, it can only be achieved if learners, interpreted here as members of a specific literate community, are invested in a writing system, which requires that they accept it. Indeed, in most of the literature on the topic, it is emphasized how important it is to involve the local community and that it can have detrimental effects if specialists who devise a new writing system fail to do so (cf. also Karan 2014: 132).[257]

256 I will focus on the creation of writing systems in general and not on the creation of writing systems for endangered languages, an even more complex situation that requires careful consideration of many additional factors (cf. Jones & Mooney 2017).

257 In some cases, people from the 'outside' made decisions without consulting members of the community who were mainly affected by these decisions. A striking example is the *Rejaf Language Conference* held in 1928 in the Sudan, where linguistic and orthographic matters were discussed with "'native voices' conspicuously absent"

Whether a community supports the externally-controlled efforts of introducing literacy must be assessed by checking whether a number of criteria are met:

> (i) the usefulness of a literacy programme must be recognized and approved by traditional community members (e.g., elders, politicians, religious leaders); (ii) local contexts for literacy must be identified and approved by community members; (iii) there must be continued widespread use of the […] language; (iv) there must be support for the maintenance of local literacy by (local) educational systems. (Jones & Mooney 2017: 6)

An additional criterion is the consideration of pre-existing writing systems, also referred to as *legacy orthographies* (cf. Jones & Mooney 2017: 30): If a writing system already exists in a linguistic community, it is imperative to take it into account. This includes a careful assessment of questions such as: Who uses the legacy orthography? How accepted is it within its community? Notably, in such a scenario, the main task of literacy development shifts from the creation of a new writing system to, basically, the reform of an existing one.

Now, the development of a writing system and, on a larger scale, the implementation of a literacy program should only proceed if all the criteria mentioned by Jones & Mooney are met. In this context, one way of gaining "native-speaker input" (Jones & Mooney 2017: 24) is by forming committees with members from all dialects or varieties of a language who, together with linguists, should be invited to work on creating a writing system (cf. Grenoble & Whaley 2006: 156). This, of course, can prove a challenging task if no members of the community are literate in any language and have not had prior access to formal education (cf. Cahill 2014: 22).

Before discussing Unseth's (2005) criteria, it is noteworthy that the first two of them – identification with vs. distanciation from a different group – are based on the same global parameter of sociocultural naturalness: *indexicality*. While the indexicality discussed in the context of the linguistic fit in Section 6.3 was endophoric, the indexicality referred to here is of exophoric nature. In this context, 'exophoric' means that writing systems, both with their materiality and their linguistic and communicative functions, can index a multitude of social variables that lie outside of themselves and the language they are based on. To allow for a more fine-grained analysis, it is useful to work with two subtypes: *intrasystemic exophoric indexicality* and *intersystemic exophoric indexicality*. The examples discussed below highlight communities' wishes to express that they belong or do not belong to other groups. As they involve the relation between two writing systems, they are of the intersystemic type. In a later portion of this section, intrasystemic exophoric indexicality will also be addressed. It is crucially dependent on grapholinguistic variation and focused on how certain features within one writing system can be indexical.

---

(Abdelhay, Makoni & Makoni 2018: 100). Another example is literacy development programs in Micronesia in the 1970s, the outcome of which was not accepted by community members (cf. Yunick 2000; Rehg 2004).

The writing systems of Arabic and Chinese serve as fitting examples of Unseth's first factor, (1) *identification with a group*. In both languages (or groups of languages), we find a diglossic situation. In Arabic, although there are many spoken dialects, the written language is the same for all: *Modern Standard Arabic* (MSA). This poses a challenge for children in literacy acquisition, as the language they learn to write often differs in significant respects from the language they speak (cf. Saiegh-Haddad & Spolsky 2014). Socioculturally, however, a unified writing system – and this is not reduced to Arabic script with its characteristic visual appearance – unites the Arabic realm. This, crucially, has intricate cultural and religious consequences. Chinese, too, is not a single language, but a collective term for several spoken varieties:[258] While some linguists classify these varieties as dialects, others classify them as separate languages on the basis of their mutual unintelligibility (cf. Chen 2004). The choice of treating them as dialects of one language is politically motivated. It is enforced by the fact that all varieties use the same writing system and that, since that writing system is morphographic, users of mutually unintelligible spoken varieties can communicate with each other by means of the written modality.

As evidenced by these examples, the Arabic and Chinese writing systems are tools for political and cultural unification. Note that in both cases, however, some spoken varieties are advantaged: A point that is often discussed under the heading of "dialects" in the course of orthography development is the fact that such a unified writing system, employed by a range of spoken linguistic varieties, must inevitably favor some varieties (whether full-fledged languages or dialects) over others. For the speakers of some Arabic varieties, the diglossic situation constituted by the linguistic distance between their spoken variety and the formal written standard will not be as challenging as for other varieties since their variety is simply linguistically closer to the standard. The situation in Chinese is similar: Here, a Beijing dialect of Mandarin serves as the standard. Consequently, children with this specific dialect as their L1 are advantaged in acquiring the writing system while other children (and generally users) are disadvantaged both in the acquisition and use of the writing system. As Li (2018: 149) puts it, "speakers of Chinese 'dialects' such as Cantonese do not have the benefit of 'writing as one speaks'" (see below).

The preceding examples were concerned with the unity of one cultural or political and, in turn, linguistic whole. However, the wish of identifying with a certain group through the writing system frequently concerns hierarchically asymmetrical situations that involve a dominant community (with a dominant language and writing system) and a "subordinate" community that seeks to show

258 For this reason, the practice of referring to the cluster of varieties as 'Chinese' that is adhered to also in this book is not unproblematic. However, here, it is ironically also fitting since this is a study about writing and it is precisely the Chinese writing system that unites the different spoken varieties of Chinese. Notably, though, the phonological representations given for Chinese morphemes are taken exclusively from Mandarin.

affiliation with the dominant one. Thus, the desire of signaling belonging is often a cross-cultural phenomenon in which the influence of other (especially geographically neighboring) cultures, countries, and languages as well as their writing systems is of utmost importance (see below for the relevance of contact phenomena). Of special interest in this respect is Smalley's criterion of (II) maximum transfer. It helps evaluate how easily the knowledge acquired in learning to read and write a newly created writing system can be put to use when acquiring a second writing system, which in Smalley's sense is the "dominant" writing system that is used to participate in developments on a larger scale (see (3) below). "Knowledge", as used here, pertains predominantly to the graphematic module, i.e. competence with respect to graphematic relations.

To give an example: In a study that highlights processes of graphematic transfer, Hinton (2014) differentiates between so-called *English-based Practical Orthographies* (EPOs) and *Linguistic Practical Orthographies* (LPOs). She observes that linguists entering a community from the outside in order to devise a writing system for and with a community often prefer "code-internal" design factors (cf. Hinton 2014: 139). In other words, in the way they design a writing system, they prioritize the linguistic fit (and, through this detour, the processing fit) and prefer LPOs, i.e. writing systems that are linguistically custom-tailored for a language. However, the communities themselves often favor "code-external considerations". For instance, the communities that Hinton mentions are heavily biased by their familiarity with the writing system of English. This results in a preference of EPOs and leads to a situation in which at least some of the graphematic relations of the English writing system are adopted by the new system despite the fact that they interfere with the linguistic fit, i.e. are not suitable for the language that is to be written. Consider Hinton's (2014: 142) description of programs that strive to revitalize American Indian languages:

> In a growing number of cases, the community leaders of the language programs thus developed are dominant in English, and being highly literate in English, they may insist on utilizing English spelling rules in the writing systems for their heritage languages. These English-based Practical Orthographies [...] are becoming more common and are sometimes even replacing already-established LPOs.

Hinton also bases her study on Smalley's (1964) five original criteria (cf. Section 4.1) and claims people "will always use ease of learning and especially maximum transfer as their highest criteria and will therefore fall back on writing systems that reflect their knowledge of the dominant language orthography" (Hinton 2014: 146). This is not the place to discuss at length how this adherence to dominant languages is problematic (especially from the perspective of *colonial linguistics*, cf. also Abdelhay, Makoni & Makoni 2018), but it is worthy to point out how striking it is that communities often wish for their writing systems to resemble different writing systems rather than to find their own 'written identity'. Smalley (1964: 36) provides this telling example:

> In Latin America many a missionary has worked out a splendidly consistent writing system based upon linguistic principles and the use of a practical phonemic alphabet. He has usually found, however, that he needed to modify that system in the direction of Spanish spelling usage, even where it introduced a limited amount of inconsistency. The influence of 'educated' bilinguals, the prestige of identification with Spanish culture, and the elements of transfer value have all united to make new literates want to learn a system as close as possible to the prestige language around them.

As described at the beginning of Part II of this book, it is theoretically possible to merely adopt a script, i.e. a (historically developed) set of visual basic shapes. However, commonly, (part of) the graphematic module of a dominant writing system is adopted along with a script. Note that the writing system itself cannot be adopted since a given writing system is always intricately linked to a specific underlying language. Yet, the graphematic module of a writing system can be partially adopted in that its graphematic relations can be "recycled": If the source and the target language share a number of the same phonemes,[259] for example, the basic shapes of the donor script that are borrowed by the target writing system can be associated with analogous linguistic units as in the donor writing system. This means that ultimately not only basic shapes but also their graphematic values are adopted: In the above-mentioned EPOs, for instance, the vowel grapheme <a> will always be used to correspond with phonemes that are similar to the ones it relates to in the donor writing system English. In this context, Smalley (1964: 45) mentions a situation in which, in his opinion, the criterion of making possible "maximum transfer" was gravely violated: in some early attempts to write tribal languages of Thailand using the writing system of Thai, both basic shapes and graphematic relations were adopted. However, one particular and very idiosyncratic feature of the Thai writing system was discarded: the non-linear arrangement of vowel graphemes that is typical of abugidas (cf. Section 6.7). Instead of positioning vowel graphemes above, below, or next to (i.e. both to the left and right of) consonant graphemes, as is usually done in Thai, the Thai-based writing systems for tribal languages "simply used the symbols but with Roman sequence of left to right" (Smalley 1964: 45). This Western-influenced linear arrangement diminished the degree of transferability of these Thai-based writing systems greatly as their readers basically knew Thai graphemes thanks to the borrowed graphematic relations but were not familiar with their spatial arrangement in Thai. Thus, their writing systems were not as useful as they potentially could have been in helping users read and write Thai, the dominant language that would allow participating in developments on a larger scale (see below).

---

**259** The same can also be the case for morphemes, of course. When basic shapes of Chinese script, which are used for morphographic graphemes, were borrowed by Japanese (as kanji), they were, in most cases, associated with Japanese morphemes that are analogous to Chinese morphemes. |木|, for instance, is thus associated with the Chinese and Japanese morphemes for *tree*, respectively. Crucially, these morphemes are distinct linguistic units of two different languages, but they have a similar status in these languages. In such a case, one can speak of the borrowing not only of a basic shape but also of a graphematic relation.

At this point, a crucial difference between original Naturalness Theory and a naturalist-based functional treatment of writing must be emphasized: when it comes to the use of scripts and, to some degree, even the basic graphematic relations of writing systems, frequency of use or occurrence cannot serve as a heuristic to evaluate what is (more) natural in writing, at least not in a straightforward manner. This is largely due to the circumstances described above: the choice of a script, and in general, the creation of a writing system, are processes that are intricately (and inseparably) linked to sociocultural factors. Take the fact that so many writing systems of the world employ Roman script. This cannot and should not be interpreted as evidence for the claim that Roman script is the most natural script (cf. Sebba 2009: 41). It is rather hegemony that is reflected in this very ubiquity of Roman script. Partially, it can be attributed to the missionaries who act as orthography mediators and propagate the use of Roman script but, of equal or even greater importance, it is also due to the communities' own wishes to associate themselves with dominant social (political, religious …) groups who use Roman script for their writing systems. These power relations affect new writing systems to varying degrees. Their graphematic influence can be conceptualized as a continuum: on the one end, there are writing systems that merely choose to use Roman script while retaining no or only few of the graphematic relations of their donor language (be it Spanish, English, French, etc.), and on the other, there are writing systems that adopt nearly all of the donor writing system's graphematic relations or at least as many as the structure and the units of their target language allow.[260]

The second factor described by Unseth and the polar opposite of wishing to identify with a group is (2) *distancing oneself from a group*, also referred to as *ideological distanciation* (cf. Jones & Mooney 2017: 25). In cultural studies, this strategy is also known as *othering* (cf. Dervin 2012). Many cases that fall into this category are characterized by communities wishing to "create ideological independence from former colonial powers" (Jones & Mooney 2017: 25). As for specific reasons motivating this wish to signal distance from a particular group, Unseth (2005: 24) mentions that it can be fueled by "ethnic pride, desire for political autonomy, religion", among others, adding also that "these categories are not mutually exclusive". All three mentioned factors – ethnicity, politics, religion – play a role in the two arguably best known examples of this strategy. The first of them is Croatian vs. Serbian. While both languages use alphabetic writing systems, the Orthodox Serbs employ Cyrillic script for their writing system while the Catholic Croats use Roman script for theirs (and the situation is even more complex, as Serbs use both scripts and their choice is based on a complex bundle of interacting factors, cf. Section 8.1). Similarly, Urdu employs Arabic script – and with it, an abjad – while Hindi uses Devanāgarī for a writing system that is typologically an

260 For a target phonographic writing system to adopt the graphematic relations of a source phonographic writing system, for example, the phoneme inventories of the two languages must exhibit some similarities, i.e. they have to partially overlap. Otherwise, graphematic relations could not be successfully 'reassociated' with analogous phonemes in the target language (see above).

abugida. Even though spoken Urdu and Hindi are mutually intelligible, "the cleavage between [them] is experienced so strongly by many speakers that they deny intercommunicability" (Coulmas 2003: 232), which the different scripts are meant to reinforce. Another example is North Korea abolishing Chinese characters (in Korean called *hancha*), while they are – even though in a restricted way – still in use in South Korea (cf. Coulmas 2016: 48–50). This abolition in North Korea was, unsurprisingly, motivated by political reasons. In extreme cases, as Unseth (2005: 25) comments, the wish to signal distance from another group is expressed by the creation of an entirely new writing system instead of the adoption of an already existing one. In many cases, this also involves the creation of a new script; in this context, it is notable that, in the course of literacy development, communities often interpret scripts *pars pro toto* as the writing systems that are visualized by them. In fact, in the creation of new scripts (as discussed in the graphetic sociocultural fit, cf. the previous section), the desire for dis-identification can be so strong that it can even affect the visual makeup of the newly created shapes, i.e. they are purposely designed in ways that make them as visually dissimilar from another script as possible (cf. Unseth 2005: 26).

Unseth (2005: 24) describes a very illustrative example of the strategy of distanciation. It concerns the Daasanach, a people who live at the Ethiopian-Kenyan border: The large majority of them (over 30,000) live in Ethiopia, while a small fraction (of only over 2,000) live in Kenya. Notably, the only literature in their language has been produced in Kenya using Roman script. On the Ethiopian side of the border, the sociocultural situation is more complex: there, Roman script is already in use by the Oromo, a people whom the Daasanach feel dominated by. This is cited as the reason underlying the Ethiopian Daasanach's decision to use the Ethiopian syllabary instead of Roman script for their writing system. This is a perfect example of system-specific sociocultural naturalness at a very small, local scale: for the Ethiopian Daasanach, the use of the Ethiopian syllabary or, more importantly, the non-use of Roman script that was already used by the Oromo is 'natural' since it signals social distance from the people by whom they feel dominated. This, as so often, goes against the linguistic and likely also the processing fits: in the Ethiopian syllabary, it is impossible to indicate vowel length, which, however, is phonemic in Daasanach. Also, Daasanach has some consonants that are not available in the Ethiopian syllabary and, thus, cannot be written (cf. Unseth 2005: 24). Furthermore, not only the linguistic and processing fits are violated by this choice but also other facets of the sociocultural fit: because of their wish to distance themselves graphematically from the Oromo, the Ethiopian Daasanach automatically also distance themselves from the Kenyan Daasanach who, like the Oromo, use Roman script. Seen from a broader perspective, they also choose a rather marginal script and writing system when it comes to participation in global affairs (see below). Adopting Roman script would have been more natural in that specific regard, since it would have allowed "maximum transfer" (in Smalley's terms), making possible the acquisition of internationally important languages and, crucially, also the use of readily available technology suited for Roman script.

China's insistence to maintain its morphographic system despite efforts to Romanize or alphabetize it is another process that can be placed into this category (cf. DeFrancis 1943). The wish to keep the morphographic system is arguably mainly due to the expression and preservation of cultural and traditional continuity. Chinese writing, as the product of one of the original independent ancient grammatogenies, is the oldest writing system still in use today. Changing it (and in such an invasive way) would not only equal an abandonment of thousands of years of culture and history but would also feed into the Eurocentric criticisms that have discarded the Chinese (and Japanese, ...) writing systems as being less efficient than alphabets (cf. Hannas 1997). This view has been largely superseded today, given the cultural prosperity of China, Japan, and other Asian cultures, and especially their economic force, high levels of literacy, and high quality of education (cf. Sampson 2016: 564).

The third major sociocultural factor that was already mentioned several times is (3) *participation in developments on a larger scale*. It boils down to the fact that the choice of a writing system "clearly influences a group's preparedness to interact with other groups outside their circle, regionally or internationally" (Unseth 2005: 27). Several subfactors must be considered here. First, the choice of script. In the creation of new writing systems, this choice is often influenced by the wish not only to be associated with a different, in most cases dominant, group (see (1) above) but also by the wish to be able to communicate with that group. The wish of being able to communicate at a more global level is also influenced by technological factors, which are subsumed by Smalley's (1964) criterion of (III) maximum ease of reproduction. Existing scripts, especially if they are used by many writing systems across the globe, usually exhibit a high level of technological implementation. Take Roman script (again), for which most modern technology is specifically tailored, beginning with (now superseded) typewriters or, more contemporarily, keyboards, input methods for mobile phones, tablets, etc. Technology is not as suitable for most other scripts, but there still exist solutions for many of them, e.g. the phonetically-based input methods for Chinese and Japanese. Even if only through an alphabetic detour, these morphographic (and syllabographic) systems can be used on most technological devices as well.[261] Many – though certainly not all[262] – scripts of the world in use today are encoded in Unicode, which means they can be written electronically and typefaces are available for them (cf. Dürscheid 2018). Hence, the importance of technological factors is "diminishing with the rapid development of computer software" (Unseth 2005: 28). In other words, while Roman script is without a doubt still the "most natural" choice with respect to the technological fit, on a continuum, many other existing scripts also fare well. By contrast, something that is categorically dispreferred is

261 On touchscreens, users can nowadays often also write characters by hand – with their fingers or (special) pens that can be used on screens.

262 The website http://www.worldswritingsystems.org/ (June 24th, 2020) features most of the world's scripts as well as information on whether they are (already) encoded in Unicode or not.

the creation of a new script from scratch. Such a script would be, at least immediately after its inception (and probably for quite some time after that, as technological implementation takes time), technologically ill-equipped. It would, thus, not only be somewhat of a visual barrier in the communication with other languages and their writing systems (see below) but also a disadvantage in a purely technological sense. On the other hand, it would of course allow designing a script (and writing system) that is intimately tied to a specific community and culture. This highlights that different parameters of the sociocultural fit are in conflict.

Once a decision concerning the visual appearance of a writing system has been made, in the next step, the possibility of graphematic transfer discussed above under (1) comes to the forefront. If some of the graphemes of a new writing system have "borrowed" graphematic relations from English, this can significantly facilitate users' spelling acquisition in the course of learning English as an L2. Since English is so dominant internationally, this puts people whose writing systems are based on the English writing system at a certain advantage.

Which conclusions can be drawn from the three sociocultural factors discussed? One obvious conclusion is that whatever a community (largely) agrees on internally counts as socioculturally natural. This can also not easily be overwritten by external orthography mediators and their rational, mostly linguistic reasoning. All three factors might converge when a community wants to distance itself from a group that oppresses it but simultaneously wishes to graphically move closer to a different (most often also dominant) group and its language. This, in turn, could even allow the community to use its new writing system to participate in developments on a larger scale – provided that the dominant language that is used as orientation is politically relevant more globally, probably even internationally as well as suited for technology. What is important for all three factors – as it also conflicts with them – is a pair of arguably the most universal naturalness parameters of the sociocultural fit: *availability* and *accessibility*. In a literate community, literacy should be available and accessible to as many members of the community as possible. This, of course, depends on a variety of prerequisites and calls for a differentiation between external and (relatively) internal factors. Factors such as education or economic status of a community's members are external: they are political factors that are, of course, very likely to affect literacy levels. However, since they are external, they cannot readily be used to assess the sociocultural naturalness of a writing system. By contrast, two factors that are internal and can ensure availability and accessibility are the choice of (a) a standard variety and (b) of a writing system whose technology is most accessible to everyone.

The (a) *choice of a standard variety* was already addressed in the context of Chinese. It concerns the fact that of the multiple varieties that may be spoken in a community, the choice of one variety to serve as the basis for the standard is influenced by a number of considerations such as: which variety is spoken by most of the members of a community? If such quantitative considerations are not helpful, e.g. when the use of varieties is distributed equally, could it be more natural to mix varieties so that the written standard incorporates certain features from all of them? Or does that make the picture too chaotic and the written standard consequently more difficult to acquire for everyone as it does not correspond perfectly with a single variety? Crucially, this choice of a (spoken) basis for a writing system is one of the sites of ideological interference. Here, prestige plays a dominant role. Decisions in this context are often not based on what is more natural for most people but on what is most convenient politically for one group within a community. Such is the case with the choice of the Beijing dialect as the basis for literacy education in China (cf. Li 2018). Given that the choice of a standard variety is, as already evidenced by its name, a form of *standardization* and, in most cases, a form of external regulation, it will be discussed in more detail in the context of the naturalness of orthography (cf. Section 8.3).

(b) *Technological accessibility* is arguably both an external and internal factor – depending on the perspective taken. Given the choice, a script and writing system should be chosen or devised for which technology is already available; this, as mentioned above, includes keyboards, typefaces, software, etc. In other words, the choice should not hinder a community from making use of these technological communicative tools. In this context, Jones & Mooney (2017: 31) specifically warn that "non-Unicode characters" should "not be used [...] given the ever-present and growing use of word processing" and that "text messaging is often found to be a predominant context for writing, and so only characters that are replicable on mobile phones should be used in these contexts". These technological factors are internal insofar as there exist some scripts (the most extreme example of which is of course Roman script) for which a lot of technology is available. This, although it has nothing to do with the script *per se*, is still intricately linked to the script. The general availability of technological devices and possibilities for a given script has, to repeat a sentiment from above, much more to do with hegemony and politics rather than an inherent suitability of a script for technology.[263] The perfect technological choice, thus, is inherently politically biased. From a different perspective, technology, much like education and economical status mentioned above, is external, for even if technology is in theory available for a script, this does not automatically mean it is available in every community. For instance, if a community lacks access to the internet let alone computers or other electronic devices, the possibilities provided by scripts and writing systems that are well-implemented technologically are severely curtailed. This leads to another important point: If the technological suitability of an existing script that makes it an attractive candidate for a new writing system is counterbalanced by the lack of available technological equipment in a community, this might render a different choice more natural. Indeed, if 'it does not matter' technologically whether, for example, Roman script is chosen, this opens up the possibility to create a new script with basic shapes that have natural systematic and processing fits on the one hand and – crucially –

263 However, some features of scripts might influence their inherent suitability for technology. For example, the size of scripts – defined as the number of basic shapes – posed, in the case of Chinese, a challenge for the design of a Chinese typewriter (cf. Mullaney 2017).

are perfectly suited for a given community as they are created specifically by/for them. In any case, these considerations are largely system-specific and must thus be assessed individually for each new writing system to be created.

Let us now turn from sociocultural factors relevant in the creation of new writing systems to those central in existing and established writing systems and literate communities. This is where the second type of the above-mentioned sociocultural indexicality comes into play: intrasystemic exophoric indexicality. It refers to the fact that certain features of writing can function indexically; this allows not only entire literate communities but also, at a lower level, their individual members to position and present themselves socially (cf. Spitzmüller 2013, 2016, in press).

The cornerstone of exophoric indexicality or, more generally, the potential of social meaning that is conveyed beyond (or in addition to) merely linguistic meaning is grapholinguistic variation. It is at the core of conflicts of the sociocultural fit with both the linguistic and processing fits. As was elaborated in earlier sections of this book (e.g. Section 2.3), variation – at both the graphetic and graphematic levels – amplifies the graphetic and graphematic solution spaces, respectively. Graphematically, for instance, it reduces transparency and uniformity. This can result in difficulties in the processing of writing systems that exhibit a high degree of variation. However, from a sociocultural perspective, variation is an invaluable resource. Indeed, it can be seen as a prerequisite for interpreting linguistic behavior as social practice (cf. Spitzmüller in press). Only in writing systems in which multiple spellings (or, at the graphetic level, "designs") are possible in producing a given utterance can users fully realize writing's sociosemiotic potential. In this context, Sebba (2007: 32) notes that if writing "is to carry social meaning, then there must be scope for variation". As established in Chapter 3, in orthographically regulated writing systems, both licensed and unlicensed variation can only occur if the systems are not fully biunique. At the graphetic level, variation is likely more prevalent in scripts with basic shapes that are visually more flexible, i.e. have larger graphetic solution spaces. Returning to the conflicts between the fits, a central question is why the possibility of variation is regarded as socioculturally natural when it introduces complexity both descriptively and from a processing point of view.

To recap, without graphetic or graphematic variation, it would not be possible to convey additional social meaning at these levels of writing, i.e. meaning that extends beyond mere denotation. Imagine a scenario in which there exists only a single typeface. This is not an entirely unrealistic scenario since a few decades ago, when typewriters were heavily used, at least at first, only one typeface was available, i.e. the appearance of type was very similar even across different models of typewriters. In typewriting, there was simply no other typeface available, which means users of typewriters could not convey additional, non-denotational information about themselves or the text by means of graphetic, i.e. visual variation. Entire subbranches of design – in typography, marketing, etc. – would not work without this very variation since they rely on the many facets of meaning it potentially evokes – provided producers and addressees share the same knowledge about common uses of graphetic variation and their respective meanings (cf. Spitzmüller in press). Similarly, at the graphematic level, imagine a wholly biunique writing system in which graphematic solution spaces for words are actually redundant given that there is always only one possible spelling. In such a system, variation would not exist. If such a system were additionally regulated orthographically, i.e. the only possible spellings were codified as orthographically correct, what this orthography would lack, due to the non-existent graphematic variation, is the possibility of meaningful deviance. Consider, for example, the replacement of Spanish <c> or <qu> with <k>, a grapheme that is not part of the native grapheme inventory of Spanish but included in the grapheme inventory of the neighboring writing system of Basque and many other European writing systems. This replacement is a practice of mainstream text messaging in Spanish, has its roots in Spanish counterculture, and, before coming into use in digital texting, was popular in graffiti during the late 1990s (cf. Sebba 2009: 43). It 'works' only because there is more than one way of representing /k/, in this case not in the graphematic module of Spanish but in graphematic modules of other writing systems that are known to many users of Spanish. This example is not only a case of variation but more specifically of unlicensed variation. Deviating from the standard carries meaning, but so does the connection that can be established between Spanish and neighboring writing systems such as Basque, which evokes a whole slate of additional facets of meaning (e.g. political meaning).

Whether stripping users of the possibility to socially express themselves through variation in writing is unnatural or, inversely, just how much variation in a writing system is natural are, ultimately, philosophical questions that cannot be answered empirically. What I want to argue here is that variation fulfills, among other things, crucial social functions. And if, very simplistically, we conceptualize the possibility of variation as an absolute dichotomy of variation vs. non-variation, then writing systems that allow for variation are more natural than writing systems that do not.

Another phenomenon that is relevant from a sociocultural point of view is the "death" of scripts and writing systems. Concerning this issue, Watt (1981: 306) claims that scripts die because of their "markedness" or, more specifically – if his views are reinterpreted within the framework of Naturalness Theory – cognitive unnaturalness (cf. also Morin 2018: 666). This highlights the importance of the processing fit. By contrast, I want to argue that while cognitive unnaturalness might be a major factor in the diachronic development of scripts, in the process of which unnatural features are changed, it is not the most important driving force behind the death, i.e. the end of existence of scripts (as well as writing systems). Script and writing system death is predominantly caused by sociocultural factors. 'Death', here, must be terminologically clarified: scripts and writing systems do not actually cease to exist, they only fall out of use. Therefore, Houston, Baines & Cooper (2003) prefer to term this phenomenon 'script obsolescence' instead of script death, a suggestion I will take up here. For script obsolescence to occur, now, special circumstances must arise. As Unseth (2005: 30) notes, literate communities usually not just abandon literacy[264] but instead switch from one writing system to a different one, with the latter being "usually from a neighboring group". It is through this switch that the initial script or writing system ceases to be used. The final stages in the use of a script or writing system are reached when only few users are left – which was the case for Nüshu women's script, Ban Niang Wu, and Dongba, for example – and/or when their use is restricted to very few, and also mostly specialized, domains. One example of the latter scenario is the Indic-derived script used to write Hanunó'o in the Philippines. Today, it is predominantly used for love songs and courtship. Although this domain is entrenched with cultural value, "it does not seem likely that a script with such a restricted use will survive the encroachment of the Roman alphabet" (Unseth 2005: 30) that is used to write Hanunó'o in all other domains.

Socially initiated change of this kind cannot be predicted. Indeed, there is no way of prognosticating when, for example, a political leader or government will decide that a writing system's script is changed and/or officially forbid the use of a previously used script (as in the case of Arabic for writing Azeri, cf. Hatcher 2008; Section 8.1). These events potentially lead to the obsolescence of the replaced and/or prohibited script. In short, the mentioned symptoms of restricted use for few (and specialized) contexts and a small number of users who know and use a script might be indicators of an approaching script obsolescence, but ultimately they are symptoms of unpredictable human behavior instead of reflexes of human physiological or cognitive pressure. While the change of scripts and writing systems can, to some degree, be predicted by their systematic, linguistic, and processing fits, the sociocultural fit adds a large portion of unpredictability to the mix. Effectively, sociocultural naturalness can only point to possibilities of a script or writing system being used, for many of the reasons named above: it is accessible to everyone, it is socioculturally suitable because it conveys belonging to or distance from certain groups, etc.

An interesting facet of the graphematic sociocultural fit that needs to be mentioned (though only cursorily) is the question of how well writing systems cope with depicting reality (or, to take the inverse perspective and invoke linguistic relativity, how they shape its users' reality), and especially how well they reflect the specifics of a given culture or society. One of many aspects that have come under linguistic scrutiny again is the status of genders, and particularly the status of women in a culture, and the way this status is reflected in and by language. A striking grapholinguistic example of this is the use of the female radical, i.e. the semantic component that signals "female" in the Chinese writing system as it is based on the grapheme <女> *nǚ* 'woman'. In their seemingly little-received analysis, Chin & Burridge (1993: 54) establish that 90% of words that include the female

264 Unseth (2005: 31) mentions cases in which literacy was indeed 'lost': Rongorongo script on Easter Island, Indus (Valley) script in India, and the loss of writing in Greece during the Dorian age.

radical "are either semantically negative or convey images of women steeped in damaging stereotypes" and conclude, to put it rather drastically, that the Chinese writing system is sexist in its portrayal of women.[265] A demonstrative example of this is <妇> *fù* 'wife/married woman', made up of the female radical and an additional semantic component with the meaning 'to sweep'. This gives the impression that a married woman is a "woman who wields the broom" (Chin & Burridge 1993: 64). Other examples include reduplications of the female radical as in <奻> *nuán* 'quarrel, dispute', in which it occurs twice, or <姦> *jiān* 'adultery', which includes three instances of the radical. One can argue that the structure of these graphemes reflects historical societal beliefs and structures that were predominant at some point in history (or still are, to some degree), and I do not want to get into the specifics of the contemporary problematics of these historic remnants in the modern writing system of Chinese. Instead, these examples are intended to simply highlight that there is a wealth of dimensions to writing, many of which are fundamentally socioculturally charged. Chin & Burridge (1993) point out that there are generally a number of facets in Chinese – predominantly its lexicon – that they consider sexist, but writing 'takes the cake' in that the degree of sexism it exhibits is vast and, importantly, largely independent of the sexism that is part of the other levels of the language. Writing, thus, can be indexical in ways that are not linked to other levels of language, highlighting once again that an independent and elaborate study of writing and its idiosyncratic systematics is a worthy endeavor.

Finally, this section cannot conclude without an explicit mention of contact phenomena. From a sociolinguistic perspective, language contact is a central issue, and this is certainly also true for contact between writing systems. For example, all of the strategies described by Unseth (2005) and discussed above – wishing to signal (1) affiliation, (2) distanciation or to (3) participate in developments on a larger scale – can only work when communities are aware of each other and each other's literacy. It is obvious that existing writing systems often heavily influence the process of literacy development in neighboring communities. This highlights the fact that contact between literate communities and, crucially, also contact between non-literate and literate communities, is often characterized by hierarchical structures, most commonly by an asymmetry of power and prestige (cf. Coulmas 2014). Today, writing systems do not exist in isolation, and in times of globalization and the advent of digital communication, the ways that literacy in a given community might be affected by external factors have only multiplied. A detailed discussion of contact phenomena would exceed the scope of this book, but it should nonetheless be clearly stated that contact must always be considered as

265 Interestingly, in a study in which 43 college students from Taiwan were instructed to rate 323 characters that include a semantic component (either the female radical, the son radical, or the human radical), the authors found that while characters with the son radical were more positively rated than those that include the female radical, the latter were also evaluated positively. They conclude that "gender inequality does not find itself in the gender-based characters" (Cherng, Chang & Chen 2009: 427).

a factor (or a collection of factors) that can significantly affect a writing system's sociocultural fit.

## 8.3 Orthography

> It is, of course, an open question as to how much, if any, standardization is really required. (Trudgill 2000: 136)

> [...] it may be best to allow a standard to evolve naturally instead of prescribing right from the start how a given language should be written. (Karan 2014: 109)

Orthography is different. Firstly, as established in the descriptive treatment of orthography in Chapter 3, unlike the graphetic and graphematic modules, the orthographic module is optional – writing systems do not necessarily have it. This does not mean, however, that orthography is merely a detached apparatus of rules that is superimposed upon an unstandardized writing system. To stress this again, as Schmidt (2018) argues, in writing systems that have an orthographic module, it is actually phenomenologically primary, while the graphematic module is secondary. This is reflected, for instance, in the fact that children acquiring such writing systems learn not only to write but to write *correctly* – and they do so right from the outset. This results in a situation in which behavior that falls outside of the norms of orthographic standardization always leaves the aftertaste of being a deviation. Accordingly, the resources that a writing system offers are – as criticized by Maas (2015) – severely curtailed by an orthography. This affects users in profound ways: Orthographic norms might be external at first, existing as codified rules in rulebooks and dictionaries, but due to the (socially) binding status afforded to such norms in literate communities, users internalize them. Crucially, disentangling external from internal norms and distinguishing their respective effects on the written output appears to be an almost impossible task in the context of grapholinguistic research.

There are two main approaches to investigating naturalness at the orthographic level. The first of them is rather static: Describing the status quo by analyzing a writing system's orthographic module, its rules, its exceptions, etc., and investigating its relationship with the other modules: predominantly graphematics but also graphetics. Relations between codified orthographic rules and the corresponding systematics of the graphematic module can be treated analogously to graphematic relations between visual units and linguistic units – as semiotic relations. Take the parameter of transparency (cf. Section 6.4). It can be used also to assess how the orthographic module fits the graphematic module: How transparent is an orthographic rule with respect to the aspects of the graphematic module that it regulates? The same can be done for orthographic rules that concern the graphetic module: What is the relation between what the graphetic module naturally "offers" as resources and how the orthographic module subjects it to regulation?

This first, static approach is important in the evaluation of the descriptive naturalness of an orthography and the discovery of unnatural features. These are mostly part of what I termed *unsystematic orthography*; in other words, they are rules that are not based on the existing regularities of the writing system (cf. Section 3.2). In the terms of the distinction between the systematic, linguistic, processing, and sociocultural fits, this static and descriptive view could be categorized as the semiotic fit of an orthography with respect to the other modules of a writing system. The second approach, which this section is devoted to, is by comparison rather dynamic. It can be classified (albeit not neatly) as the sociocultural fit of orthography.

The central aspect of interest for a functional usage-based investigation of orthography is arguably how the agents who decide on orthographies (whoever they might be in a given literate community) behave. This subsumes their responsibilities and the guiding maxims that drive their decision-making both in the initial establishment of an orthography and the reform of an existing orthography. No matter how internalized and primary orthographic rules become in a writing system, to some degree, they always remain an external and artificial intrusion. And while orthographies are an integral part of the system and its use, they can, in many cases, only be changed through official reforms and only by authorities who have the right to make such decisions. At first glance, this might give the impression that orthography *per se* cannot be natural. As a blanket statement, this must be rejected. For instance, a point that is often raised in favor of orthographies is that when for certain words, a multitude of possible spellings exists, licensing specific spellings as correct can facilitate communication by granting order. This is argued to help avoid 'written chaos' (cf. Grenoble & Whaley 2006: 130). However, whether such order is afforded by orthography depends crucially on the makeup of a specific orthography and, as mentioned above, on the behavior (including the motives and goals) of authorities who have the power to design it.

The first and more static of these points that concerns the interaction between the system and the norm raises the following questions: how does the orthographic module interact with the graphematic and the graphetic modules? Is the correct spelling of a word even part of the graphematic solution space? In this context, a more dynamic, usage-based question could read: is the correct spelling likely to be used or is it already in use? It would not be natural if the spelling licensed as correct by the orthography is unsystematic, i.e. not located inside the graphematic solution space and thus not part of the system (yet). It would be even more unnatural if said unsystematic spelling was additionally not employed by users; here, it must be mentioned that users do at times make use of spellings that are not established parts of the system (cf. Dürscheid 2000). Since orthographies are often not self-regulating systems but external codifications (with notable exceptions such as English), these mentioned aspects influence how authorities who make orthographic decisions can behave 'naturally'. In a nutshell, they can use their power in a reasonable way that benefits as many users of an orthography as possible. Further questions regarding the implementation of an orthography include: "[W]ho does the selecting [of a standard, D.M.]? How long might the process toward standardization take? Is standardization of the writing system a requirement for literacy to take root in a speech community?" (Karan 2014: 107f.). In some cases, as will be discussed below, the most natural action is actually not to impose a standard at all or at least to postpone its introduction to a later point. In general, two types of relevant tasks performed by "orthographers" at different stages need to be distinguished: (1) the establishment of an orthography, and (2) the modification, i.e. reform of an orthography.

The question of whether a standard is necessary in preliminary stages of a community's acquisition of literacy is discussed by Karan (2014) in the context of the development of writing systems in hitherto non-literate communities. She argues for "'slowing down' and allowing a standard to evolve through practice rather than prescription" (Karan 2014: 107) and provides examples to underline how this could be reasonable. It was mentioned before that the creation of a writing system does not equal the creation of an orthography, which is just an optional part of a writing system. Now, if the graphematic module of the newly created writing system is not completely uniform and transparent, there will be some variation in the graphematic solution space, variation that – without orthographic regulation – can be (creatively) exploited by users. In the context of literacy development, at this stage, the writing system, characterized by simultaneously occurring competing spellings, is sometimes referred to as *working orthography* (cf. Bauernschmidt 1980; Karan 2014: 108). Note that this is precisely where my criticism of the misuse of the terms *orthography* and *orthography development* becomes most pressing: the creation of a new writing system should indeed, in the first step, entail only the creation of a writing system, not an orthography. Thus, what is really done in such scenarios – or what should be done – is *writing system development* (cf. Karan 2014: 109). Admittedly, the predominant use of *orthography development* does reflect reality to some degree, since in literacy programs, standardization is often a high priority right from the start. What guides the entire process is a "'normative' expectation" (Karan 2014: 109) from the people working on creating the writing system as well as the people for whom it is created. Consequently, even in the early stages of these systems, creativity and freedom of spelling are often discouraged. As was discussed in the context of the orthographic module's functions, once literate prescriptivism has instilled in people a sense of normativity, many of them prefer not to have a choice but want clear and unambiguous rules that tell them exactly how to write instead. This trend towards standardization is welcomed (top-down) by governments who often instruct committees or agencies with the task of designing an orthography; crucially, from a political perspective, an orthography allows the production of consistently and uniformly spelled reference works – textbooks for education, dictionaries, etc. – which of course also involves economic considerations (cf. Karan 2014: 110). In short, an orthography brings order to a situation that is, through a normative lens, perceived as chaos.

This mindset is highly influenced by the contemporary relevance of orthographies. As will be discussed below, many writing systems that have existed for a long period did not initially come into existence as standardized systems.

Thus, they, too, existed without orthographies after being 'created' and gradually coming into use. Bearing this in mind, it may truly be reasonable to enforce "non-standardization" (Karan 2014: 111) in modern cases of writing system development. This, notably, does not imply a neglect to standardize due to standardization not being on the agenda of orthography mediators. Instead, it is their deliberate and conscious decision *not* to standardize a system, a decision to wait and let a natural play of forces sort out some of the graphematic conflicts. By comparison, in the case of an often rushed implementation of an artificial standardization, these conflicts would need to be solved with arbitrary, external, and top-down solutions.

Another aspect that speaks against the implementation of an orthography at too early a stage is that later modifications (which might become necessary) are a difficult affair. While modifications of the graphematic module occur naturally through use, modifications of orthographies (almost) always constitute reforms (cf. Section 3.4) and come with an entire machinery of sociopolitical consequences. Since most orthographies are external codifications, reforms represent amendments of these codifications. This has major repercussions for literate communities that have grown accustomed to their orthographies: for example, people who are competent in the superseded orthography must acquire new rules, which (at least in a transition period from the old to the new orthography) can affect the processing fit. A technological aspect that could occasion costs is the update and replacement of material using the old orthography, especially dictionaries and textbooks. These considerations underline that deciding on an orthography too early in the development of a writing system means risking the necessity of reforms that themselves entail a number of negative consequences.

But how can a natural orthography be achieved in a writing system? Karan (2014: 113) lists a few steps that can be followed:

> Writers' actual, uninhibited writing before being "trained" needs to be analyzed. Deviations from the proposed standard might reveal mother-tongue-speaker intuitions possibly indicating a point of linguistic misanalysis or highlight certain symbolization preferences.

Crucially, some literate communities may not value standardization highly. Examples detailed in the literature include Jamaican Creole English and Yucatec Maya. As Brody (2004) has shown, users of the Yucatec Maya writing system tolerate variation in spelling. In fact, a meeting in 1984 that had the purpose of establishing a written standard came under a lot of scrutiny as the public interpreted it as an exclusive enterprise "by linguists, for linguists" (Brody 2004: 155) in which native speaker input was unwelcome – input that is, however, crucial.

Karan's (2014: 119) major argument rests on the fact that "[s]tandardization and the implementation of a reform take time". She highlights this by contrasting the immediate push for standardization in modern creations of writing systems with several European writing systems in which standards developed *naturally,* i.e. over a long span of time and without (or with only little) prescriptive intervention. Elmentaler (2018) outlines major aspects of such a development in his historical account of the German writing system. Notably, while variation did not pose a problem, certain factors in the development and use of the writing system did exhibit standardizing tendencies; however, they were not perceived as interfering with the system from the outside. The German writing system was thus standardized in a long 'natural' development and only then was the resulting standardization externally codified as an official orthography. In a nutshell, this natural emergence of a standard differs fundamentally from the immediate standardization of a new writing system that has just been created from scratch.

After the decision to standardize has been made, one of the central and simultaneously socioculturally most complex factors that can "emerge as a contentious and divisive issue" (Jones & Mooney 2017: 7) is the choice of which variety of a language should serve as a basis for the standard. There are various approaches to this problem. In the (a) *unilectal approach*, a single reference variety is chosen as the basis for an orthography. One advantage of this approach is that – as in the case of Arabic or Chinese – it can unify and highlight a common identity, which can prove socioculturally natural. However, users of other varieties than the one chosen as the base variety are put at a disadvantage. They must effectively adhere to another variety. If they are not yet familiar with that variety, they must even learn it in order to be able to write correctly, i.e. bridge the linguistic distance between the variety they speak and the variety that is written. Unsurprisingly, the unilectal approach may foster social and political tensions as the choice of one variety could be interpreted as "deliberate favoritism" (Karan 2014: 116). There are several factors that influence the choice of a variety: for example, a variety might be chosen because it is spoken in a region in which political power is most highly concentrated. Furthermore, the number of a variety's speakers might also prove decisive. A third relevant factor is prestige, and often the most prestigious variety is the most likely candidate for a standard. Should authorities in charge of an orthography indeed proceed with the unilectal approach, it is paramount they make "clear to all [users of an orthography, D.M.] that adherence to a written standard allows text to be read with various dialect pronunciations" (Karan 2014: 116), as this could increase the acceptance of the unilectal orthography by users of other varieties.

In the (b) *dialectal approach*, multiple orthographies are established based on multiple varieties. This strategy may be chosen when the linguistic distance between the varieties of a language is too great. However, given the sociocultural indexicality of writing, the existence of distinct orthographies conveys social fragmentation. This can, of course, be a desired outcome in cases in which users of different varieties actually wish to distance themselves from one another, but it is detrimental in cases in which the purpose of an orthography is to highlight the unity of a language and linguistic community by means of its written standard. It must be noted that the dialectal approach is seldom the one adhered to in practice (cf. Jones & Mooney 2017: 8).

In the (c) *multi(dia)lectal* [266] *approach*, features from various varieties are combined in a single orthography. This can help "foster a common identity for the speech community at large" (Jones & Mooney 2017: 8f.). However, choosing which features of the different varieties to include in the combined orthography is no straightforward matter, and "the question of how best to accommodate different varieties within a single orthography leads directly to issues of power and authority" (Sebba 2007: 112). From a linguistic point of view, the inclusion of features from different varieties may result in situations of overrepresentation and underrepresentation: for example, when a distinctive feature in one variety that is not distinctive in another variety is included in the orthography, this leads to superfluous overrepresentation (of a non-distinctive feature) for users of the latter variety and thus intervenes with both the linguistic and processing fits (cf. Karan 2014: 117).

Another option is the (d) *common-core approach*: here, an orthography is based on an "artificial"[267] variety that is created by means of a historical reconstruction of the common core of a language's various varieties. This approach has had very limited success. It is not recommended since it is unnatural concerning its accessibility: a common-core orthography is not directly accessible to any speaker of any of a language's varieties since it is not directly based on any one of them (cf. Karan 2014: 117f.; Jones & Mooney 2017: 9). The disadvantages for all members of the language community outweigh the sociocultural benefits of this approach (e.g. that no single variety is given a privileged status).

When it comes to the choice of one variety or multiple varieties as a basis for an orthography, it is paramount for the authorities in charge to convey clearly to users that the (written) standard variety is merely an "additional variety and [...] does not aim to replace spoken dialects" (Jones & Mooney 2017: 9). Users must be reassured that the varieties that were not chosen as or included in the standard still have an equal footing. This, however, might be perceived as insincere by users when one variety is chosen since its status is elevated to that of a standard, which raises its prestige and may give the impression that the other varieties enjoy less prestige. It is, thus, debatable whether all varieties in such a situation can truly be equal.

Note that the approaches listed above are restricted to phonographic writing systems or phonographic elements in other types of writing systems. Morphographic systems pose different and unique challenges for standardization. In principle, for example, in Chinese, the "writing system permits the imposition of different phonological systems onto the syntax and the lexicon of the standard written language" (Handel 2013: 24) since, holistically, Chinese graphemes refer to

266 Various terms refer to this approach. While Jones & Mooney (2017: 8) call it *multilectal*, Karan (2014: 117) refers to it as *multidialectal*.

267 The dialect would be artificial since it is created artificially. However, it contains and combines features from the existing varieties or historical predecessors of those varieties and thus exhibits natural features. Such an artificial dialect is thus comparable to writing: it is phylogenetically artificial but contains natural features.

morphemes (whatever pronunciation these might have in a given variety). Whether the phonological "clues" provided by the phonological components of Chinese graphemes are useful for other varieties than Mandarin is a reasonable question, however. In general, in the course of the creation of new writing systems, the possibility of creating a morphographic system is often quickly discarded in favor of an alphabetic (or, though rather seldom, syllabographic) system. The reason for this might be that "it is argued that morphographic systems are inferior to phonographic ones" (Jones & Mooney 2017: 13). This, however, in this absolute way (and without a specific context) can in no way be regarded as an accurate statement, as the suitability of a type of writing system is crucially dependent on the properties of the language that it is intended for (cf. Section 6.1). With respect to morphography, it is also noteworthy that predominantly phonographic systems can (and do) incorporate morphographic features. In this vein, modern accounts argue, for example, that morphography plays a central role in the alphabetic writing system of German (cf. Schmidt 2018).

In the following, I want to discuss examples of how, from a sociocommunicative and sociocultural perspective, the change of orthographic regulation – i.e. orthography reforms (cf. also Section 3.4) – interacts with the different modules of writing systems: first the graphetic module, then – through the question of variation – the graphematic module.

An example of how orthography can interfere with the graphetic module is the addition of a basic shape to the version of Roman script used in German: in 2017, the *Council for German Orthography* added an uppercase version of |ß|, namely |ẞ|. This inclusion also has (admittedly minor) consequences for the graphematic module, as it is now possible, when writing in all caps, to use <ẞ> instead of the alternative <SS>. It also affects the graphetic module, as graphetic solution spaces shift because the original lowercase and the new uppercase basic shapes are visually very similar. Furthermore, the similarity between |ẞ| and |B| also strikes the eye. Unsurprisingly, several reactions in the media harshly criticized these visual similarities.[268] It is interesting to note that what appears to concern people the most in the discussion surrounding this addition is indeed the graphetic aspect, although its general usefulness was also questioned by many.

268 Cf. the criticisms by Andreas Platthaus (published June 30th, 2017) in the *Frankfurter Allgemeine*: http://www.faz.net/-gqz-8zbt2 (accessed July 27th, 2020) and by Cornelia Geißler (published June 29th, 2017) in the *Berliner Zeitung*: https://www.berliner-zeitung.de/kultur/neureglung-rat-fuer-deutsche-rechtschreibung-fuehrt-grossesss-ein-27883300 (accessed July 27th, 2020).

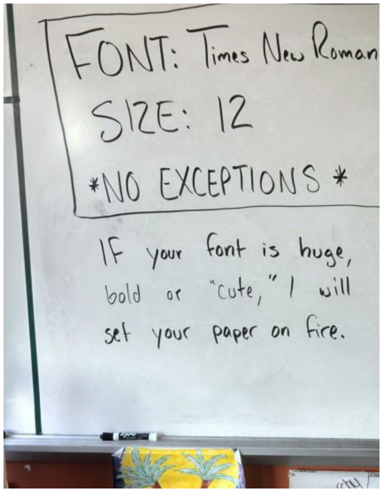



FIGURE 48. Threat of sanctions for violations of graphetic rules, from: https://i.pinimg.com/originals/86/5c/9a/865c9a6a72f44777eb4eedad9a3ee045.png (February 9th, 2019; page no longer available)

Another example of graphetic norms is the regulation of how digitally produced text is designed. For example, it is not uncommon for teachers both at schools and universities to hand out guidelines detailing how students should format papers. These guidelines operate at a local level that could be termed the micro-level: they are often specific to a given person (as an orthographic authority) or institution (such as a university department), a given context (such as a class), and a specific task (such as the design of research papers or theses). Given their specificity, the question of how strictly deviations from these norms are sanctioned also depends on the orthographic authority in charge. Some might sanction deviations strictly, announcing not to grade or even read papers that do not conform to the norms. Consider, in this context, Figure 48, a demonstrative example in which a teacher deals humoristically with the concept of orthographic power. So-called *house rules* of different publishers or publications such as newspapers can also be conceived of as a type of micro-level orthographies (cf. Schimmel-Fijalkowytsch 2018: Chapter 7).

While the rules named above are explicit, other orthographic regulations of graphetic aspects of writing might remain implicit: in a study on online discourses surrounding the (use of) typeface Comic Sans, an example illustrates how a person who identified themselves as an HR manager claimed they would immediately throw out any CVs that used Comic Sans in their designs (cf. Meletis 2020a). This reaction is based on the fact that in the same vein as ascriptions that are based on orthographic errors (see below), people tend to ascribe certain features such as informality, childishness, etc. to Comic Sans and, in turn, also to the person choosing to use this typeface. Implicit rules of this kind are arguably not part of the orthographic module *per se*, which was defined as being externally codified and binding. However, the example of Comic Sans definitely underlines the fact that the countless choices that are offered by licensed graphetic variation – e.g. in the form of thousands of typefaces to choose from – are not necessarily free but constrained by the fact that they are indexical (which users may or may not be aware of). They are, thus, subject to implicit norms and possibly even consequences that ensue if said norms are not conformed to.

These examples underline that there are ways in which the orthographic module can interact with the graphetic module. The following will deal with another core aspect: how orthography deals with variation in the graphematic module and even the orthographic module itself.

One of the features of orthographies presented in Section 3.1 is their small degree of variability. While the graphematic solution space might offer countless possibilities for spelling a specific utterance, these are curtailed by orthography. In the discussion of the sociocultural fit of the graphematic module (see above), it was established that from a communicative point of view, a certain degree of possible variation is natural and necessary for sociocultural indexicality to even be possible. How much of this variation is then actually orthographically licensed as 'correct' is an altogether different question. The smaller the graphematic solution space, the smaller (usually) the divide between graphematics and orthography, since in that case, the orthographic module merely codifies as orthographically correct the (sometimes sole) spelling that is licensed graphematically. By contrast, when the graphematic solution space is large, the divide between what is possible graphematically and what is licensed orthographically becomes quite remarkable: in that case, there is a large degree of possible variation, most of which, however, is not orthographically licensed and is thus *unlicensed variation* (cf. Sebba 2007: 30).

In some cases, more than one spelling from inside the graphematic solution space is licensed as orthographically correct, which is precisely what introduces variability to the orthographic module. In German, for example, it is presently allowed to spell the word for *typography* as <Typographie> or <Typografie>. Both are correct, which renders this an instance of *licensed variation*. The question, now, is which type of variation is more indexical: licensed or unlicensed variation? In many cases – such as <k> when used in Spanish or <x> as in <*demnäxt> (instead of <demnächst>) 'soon', which is/was used in German fanzines (cf. Androutsopoulos 2000) – unlicensed variants carry more indexical meaning. Licensed variants, too, might transport some information such as modern vs. old-fashioned (where, for example, <Typografie> is more modern and <Typographie> is more old-fashioned), but they are not charged with connotations as much as unlicensed variants are.

What applies to the graphematic sociocultural fit also applies to the orthographic module: the needs and wishes of a literate community's members should be met as best as possible. Interestingly, in the case of orthographies, what users appear to want is not variation but clear and unambiguous rules. Thus, licensed variation appears to be undesirable to users as it not only allows but *requires* choices. In other words, writers of German (provided they are aware of the options) *must* choose whether to write <Typographie> or <Typografie>, and they do not appreciate having to make such choices (cf. Nerius 2007: 39). An interesting development in this context is the introduction of recommendations by the Duden, a German dictionary which, up until the orthography reform of 1996, was a prescriptive source of German orthography. Today, the Duden merely documents German orthography the way it has been decided on by the *Council for German Orthography*. Yet, in 2006, when the 24th edition of the Duden was published, so-called recommendations first appeared. Specifically, in cases of licensed variation, the Duden now recommended one of the possible spellings: in the above-mentioned example, <Typografie> is recommended (cf. Duden 2017: 15; 1129). These recommendations are supposed to increase the usefulness and usability of the Duden following the orthography reform of 1996 that left users in want of clarity and orientation.[269] The editors of the Duden (2017: 14, my translation) justify these recommendations as follows: "the recommendations […] shall make possible a correct and consistent orthography for all those who wish for one and do not want to make their own decisions in the choice of variants".[270] Three factors are highlighted as a basis for the Duden's recommendations, and they resemble the guiding principles of the *Council for German Orthography* that will be discussed below: (1) the actual usage of written language is observed by the Duden's editorial team, (2) the needs of readers, which call for an optimal comprehensibility of texts, shall be met, and (3) the needs of the writers, too, who wish for an easy operability of the orthography, shall be satisfied (cf. Duden 2017: 15). Interestingly, the latter two of these three factors are often in conflict, which the Duden actively acknowledges. However, in observing actual usage and preferred spellings, the editorial team can determine which one of the two principles is dominant at a given time.

Concerning the observation of how users actually use a writing system as well as the decision to postpone the implementation of certain orthographic rules, an example that deals with the depiction of the reality of life in writing (specifically the depiction of genders) is noteworthy, which also links to the question already introduced in the context of the sociocultural fit of Chinese graphemes and the female radical (cf. Section 8.2): a topic recently discussed by the *Council for German Orthography* is the question of a preferred way of writing 'genderwise' appropriately, a practice for which in German, a separate verb has been coined, *zu gendern* 'speaking/writing genderwise correctly'. In the last couple of decades, different strategies of avoiding the generic masculine – both in the singular and in the plural – have developed in German, among them the so-called *internal I* (Ger-

**269** Cf. original statements given in 2006 by a spokesman of the publishing house issuing the Duden: https://derstandard.at/2477707/Duden-gibt-erstmals-Empfehlungen (July 27th, 2020).

**270** „Die Empfehlungen […] sollen all denen eine richtige und einheitliche Rechtschreibung ermöglichen, die dies wünschen und keine eigenen Entscheidungen bei der Variantenauswahl treffen möchten."

man *Binnen-I*) as in *LehrerInnen* ‘male teachers and female teachers’ and writing out both forms as in *Lehrerinnen und Lehrer* ‘female teachers and male teachers’. Additionally, the use of elements such as an underline <_>, an asterisk <*>, or a colon <:> as in *Lehrer_innen*, *Lehrer*innen*, and *Lehrer:innen* has become increasingly popular among people who wish to write in a socially more inclusive way (cf. also Haralambous & Dichy 2019). These three latter strategies aim to include also other genders and groups of people, e.g. people from the trans* community, non-binary people, and others who do not feel included by either male or female forms. To observe and study this topic, the *Council for German Orthography* has initiated a task force/research group entitled *Geschlechtergerechte Schreibung* ‘genderwise correct spelling’, which, in November 2018, presented a report in which it discloses its decision not to codify any of the existing strategies as correct (cf. Rechtschreibrat 2018). This is justified on the grounds that these above-mentioned practices of writing inclusively constitute a ‘rather new’ phenomenon; as such, they are still being actively negotiated by literate German-language communities, which, according to the Council, becomes obvious in corpus analyses that illuminate the phenomenon. Accordingly, the Council opted not to give a clear recommendation for any strategy, claiming it does not want to interfere with the ‘natural’ way of how literate communities sort this out themselves. What the task force does formulate, however, is a slate of (very general) guiding maxims that should be followed when attempting to write inclusively (cf. Rechtschreibrat 2018: 8).

From the perspective of the sociocultural fit of orthography, the Council’s behavior, specifically its decision not to codify or recommend a strategy, can be interpreted in two ways: on the one hand, it can be viewed as a sensible decision of “‘slowing down’ and allowing a standard to evolve through practice rather than prescription” (Karan 2014: 107).[271] On the other hand, it can be interpreted as the opposite, i.e. the failure to prescribe a standard, which, in this case, equals an avoidance to take a clear stand. This latter argument is based on the fact that language not only reflects but shapes reality. The multitude of strategies available for writing genderwise appropriately produce an unclear situation for users who prefer unambiguous rules (see above) and may thus result in an avoidance of writing genderwise appropriately altogether. Such a situation is also a fertile ground for skeptics and opponents[272] whose negative attitudes, certainly not necessarily

**271** This reading of the Council’s decision is advocated in an article by Hannah Lühmann, who claims that concerning this issue, the Council puts its trust in the ‘natural development of language’ and invites the people who speak and write German to negotiate the issue themselves, cf. https://www.welt.de/kultur/plus183982304/Drittes-Geschlecht-Gendern-Vertrauen-wir-der-natuerlichen-Sprachentwicklung.html (July 27th, 2020).

**272** On March 6th, 2019, the *Verein Deutsche Sprache* (VDS), a union that counts over 36,000 members and is invested in the preservation of German as a language of culture, economy, and academia, presented an appeal titled *Schluss mit dem Gender-Unfug!* (roughly translated as ‘An end to gender nonsense!’), which, in essence, is a plea to resist the strategies of speaking and writing genderwise correctly, cf. https://vds-ev.de/gegenwartsdeutsch/gendersprache/gendersprache-unterschriften/schluss-mit-dem-but often, correlate with political views and agendas. The Council has the authority to license a standard, and the fact that it chooses not, in this case, may be interpreted as a failure to fulfill an important societal duty. It is questionable, however, whether it actually is the duty of a descriptively oriented commission to decide on such critical societal questions that undoubtedly go beyond language policy. It is paramount to note that the Council at no point argued against writing inclusively. It only has yet to argue for a 'correct' way of doing it.

With respect to the wishes of a literate community's members, which prove central for an evaluation of orthography's sociocultural fit, a series of interviews I conducted with the aim of investigating people's attitudes towards orthography must be mentioned.[273] One of the main findings was that the vast majority of interviewees perceived the process of writing – as well as written products, i.e. texts – as inseparable from orthography. Thus, the normativity associated with writing appears to be internalized. This became nowhere more evident than in one participant's metapragmatic observation: "Every time that I'm writing somewhere, I'm writing, and *therefore* it has to be correct […] because otherwise, I could just leave it"[274] (my emphasis). 'Therefore', in this quote, refers to the process of writing in general. Thus, this participant established a causal link between the process of writing and the expectation that the resulting product of writing needs to be (orthographically) correct. Writing, as a process, is thus implicitly interpreted as "writing *correctly*", which corresponds with the view that orthography is phe-

gender-unfug/# (July 27th, 2020). Unsurprisingly, this appeal has reinvigorated the public and media discourses surrounding the topic, with one side supporting the VDS' resistance and the other arguing that its views are outdated. Then, in August 2020, the *Gesellschaft für Deutsche Sprache* (GfDS) expressly advised against using the 'gender asterisk' (German *Gendersternchen*) as it does not conform to the grammar or orthography of German, cf. https://gfds.de/pressemitteilung-gendersternchen/ (August 16th, 2020).

**273** In May and June 2018, I conducted 14 interviews with eight female and six male participants aged 19 to 29. (From April to June 2019, I conducted a further seven interviews, bringing the total up to 21.) These semi-structured interviews were part of a larger research project that had at its core the analysis of the depreciative public correction of others' orthographic mistakes on the internet, a practice I termed *orthographic shaming*. Actors who engage in this behavior are commonly referred to as *grammar Nazis* or *spelling Nazis* (cf. Bahlo, Becker & Steckbauer 2016; Švelch & Sherman 2018). A core part of the project was a collection and analysis of German instances of orthographic shaming found in the comments made under public Facebook posts. In the 14 interviews, participants' attitudes towards this behavior were of interest, and participants were, among other things, given an example of orthographic shaming that they reacted to. Several interview questions also focused on general attitudes towards orthography, including: What is your view on the topic of orthography? What role does orthography play in your life? Who should be allowed to decide on orthographic matters? What is your opinion on orthography reforms, and should everyone have a say in them? How do you assess your own orthographic competence?

**274** The interviews were conducted in German. My translations here are as literal as possible.

nomenologically primary to graphematics (cf. the introduction to Chapter 3). Accordingly, anything that deviates from the norm is regarded as incorrect.

Concerning deviations, it was interesting to observe that the participants appeared to exhibit the necessary knowledge to make fine-grained distinctions between errors (i.e. deviations due to competence) and mistakes (i.e. deviations due to performance). Moreover, their attitudes towards these types of deviances were different: mistakes, mostly in the form of 'typos', were scrutinized more critically than errors. Several interviewees argued that a message full of typos gives the impression that the producer did not take their time to reread the message to eliminate the "obvious" mistakes, which would have been possible given that mistakes are a matter of performance and producers should thus be able to self-correct. If they fail to do so, this is frequently interpreted as a lack of effort, and in consequence, a lack of respect for the addressee. Another distinction the participants perceived was the distinction between errors (which are by definition unconscious) and the conscious use of unlicensed variation. Choices such as the disregard for capitalization in German and the omission of punctuation were frequently named as examples of such conscious deviations. When asked about actual errors and people who regularly make them, participants were eager to deny making any ascriptions based on these mistakes. However, the answers they gave frequently contradicted this self-assessment. For instance, one interviewee noted: "[...] usually I value when people can spell correctly because if they can't, that makes their IQ sink in my head". The passive construction in this formulation reveals that ascriptions are possibly made 'involuntarily'. Rationally, as all participants emphasized themselves, there is no reason to associate orthographic errors with low intelligence or even a low level of education. Yet, ascriptions are still being made, and while they may be involuntary, they are conscious (or were at least raised to awareness during the interviews).

This points to an aspect that is central for an assessment of how natural an orthography is: its accessibility. Orthographic errors or mistakes are not legally but socially sanctioned both in the professional and private realms. Consider CVs that abound in errors or mistakes or erroneous messages on online dating platforms or dating apps, which, if the addressee is orthographically competent and ascribes personality traits based on orthographic competence, are potential reasons to disregard someone as a potential employee or partner.

These possible sanctions for deviances from orthography highlight that it is important for everyone in a literate community to have comparable access to an orthography. In order to make that possible, the orthography should not be too unnatural (mainly from a processing perspective). In the present context, this means that the orthographic module should not increase unnaturalness in the graphetic and graphematic modules but should, instead, attempt to curb unnatural features in a way Handel's (2013) suggestion for a regularization of the Chinese writing system would have done (cf. Section 5.2).

The question to which the study's participants gave the most elaborate answers focused on authorities who have a say in the design and reform of orthographies. This echoes the fact that this topic is emotionally charged (cf. Johnson 2005). Notably, none of the participants stated that they wished to be included in the orthographic decision-making process. A number of them expressed ideas such as popular referenda in which the public could be given the possibility to participate in decision-making, or committees in which members of different groups, among them students and teachers, should work together on establishing an orthography. Interestingly, when the question circled back to who should have the final say, all interviewees agreed that only 'experts' should be authorized to make decisions. In most cases, this was justified with the fact that letting 'normal people' have the final say would risk having an orthography turn out chaotic due to irrational decisions that lack a scientific basis. However, participants also noted what experts should do: observe and investigate the use of the writing system and base their orthographic decisions on these observations. Strikingly, according to the participants, this also includes bringing to a halt developments of written language that they judged as unfavorable, e.g. certain facets of youth language. In this respect, the participants also uttered very conservative views regarding, for example, the inclusion of new words in dictionaries. Here, the Duden is actually more progressive, adding new words in every updated edition of its dictionary: in 2017, <Selfie> 'selfie', <Filterblase> 'filter bubble', <Cyberkrieg> 'cyber war', and <Emoji> 'emoji' were among these new words (Duden 2017), while in 2020, words such as <genderneutral> 'gender-neutral', <Klimakrise> 'climate crisis', and <Covid-19> were added (cf. Duden 2020). In short, the preservation of the current standard, which they regard prestigious, is among the tasks the interviewees considered as central for authorities of linguistic policy.

Inversely, participants also argued that authorities of linguistic policy should have the users' needs in mind, which was most evident in the answers to questions focusing on orthography reforms. Specifically, an orthography should only be changed in ways that respond to users' needs. Here, processing needs were foregrounded, i.e. orthography reforms should make an orthography easier to learn and use. Some of the interviewees alluded to the central naturalness conflict when they urged caution concerning the balance between writers' and readers' needs. Ultimately, however, they sided more with readers: an orthography should not become too easy to spell since that would possibly make it harder to read.

Like the challenges the Duden faces in the formulation of its recommended spellings (see above), the conflicting duties ascribed to authorities who decide on matters of orthography clearly highlight the impossible task of fulfilling all needs simultaneously. The highly intricate task they take on is (re-)shaping the writing system, particularly in a way that improves upon its user-friendly qualities rather than damaging them. This means that orthographic authorities need not only identify naturalness conflicts in a vein similar to how this sketch of a functional theory of writing does it but ultimately must make prescriptive decisions based on this analysis. This, understandably, puts a lot of weight on their shoulders, and the fact that there always exist at least some negative reactions to orthography reforms (and also initial implementations of orthographies in the creation of new writing systems) reflects how a part of the public – as well as specific stakeholders such as the media, politics, and academia – unfailingly finds faults with the decisions made. In the end, not everyone can be pleased by decisions in orthographic matters. However, here, orthography is special, since unlike in graphetics and graphematics, there are specific actors that can be held responsible for their actions by the literate community.

# IV Discussion

In attempting to prove that the construction of a theory of writing is feasible, the present study brought forth a number of diverse outputs:

— a reevaluation of theoretical grapholinguistics that aims at embedding a descriptive model of the structure of writing systems and the core units and concepts it entails into a theory that not only describes but explains the nature of writing,

— a meta-study on grapholinguistic research, showcasing different subfields of the discipline, the state of the discipline, and its many remaining desiderata that reveal themselves clearly in the lack of research on several aspects of a theory of writing,

— most importantly, a systematic analysis of the structural and psycholinguistic makeup of writing systems and their processing, and, on the basis of that, a proposed theory of writing, complete with a blueprint for comparisons of writing systems, a *tertium comparationis*.

Although they overlap considerably, these points are separate strands of a larger discussion and will be treated one by one in the following.

## Reevaluation of theoretical grapholinguistics

In his rejection of a 'structural graphemics', Daniels (1991) claimed that writing cannot be treated with the same conceptual (linguistic) tools as language and based this on the argument that writing and language differ fundamentally (cf. Section 2.2). For the same reason, he discarded grapholinguistic concepts such as the much-debated grapheme, a unit that emulates (largely) uncontroversial linguistic concepts such as the phoneme or the morpheme. The problem, now, is that in the Anglo-American research area, no alternative way of describing writing systems in a unified manner and thus no method of comparing them has been established yet. Instead, the focus of research that can be labeled as 'grapholinguistic' remains on historical reconstruction and individual descriptions of writing systems. By contrast, in the German-speaking research area, the question of how writing can be described *was* studied – with an almost sole focus on German, however, rendering many of the results inapplicable to other (non-)alphabetic writing systems. Meanwhile, a theory of writing that needs to build on a comparative foundation remained a desideratum (cf. Watt 1998).

As this book has attempted to show, writing *can* be treated with the same tools and within the same theoretical frameworks as language. If this were not the case, transferring the core tenets of Naturalness Theory to the subject of writing would have been pointless. In fact, writing arguably *must* be treated by analogy with language to make possible the description and comparison of writing systems and, most of all, the establishment of a theory of writing that offers explanations of why writing is the way it is and why it has developed in this (and no other) way. Individual, isolated descriptions of diverse writing systems that are historically and philologically oriented are also necessary enterprises, as description is undoubtedly the basis of a theory of writing. In the end, however, as excellent as minute descriptions may be, they are not automatically of value for a theory of writing; what is decisive is that they are designed in a way that enables an integration into a larger theoretical framework. This means that it is not merely descriptions but descriptions made in a uniform, comparable way that become the basis of a prospective theory of writing. For such a theory to work, a general model of writing systems complete with shared concepts and terminology is an absolute necessity.

The reason why writing can be treated like language is simple: writing is based on language, and not just that: it is dependent on language. It is a myth that writing is a system independent of language or that it is even its own language system (cf. Section 2.1). As a modality of language, writing is a part of language – by virtue of which it *is* language. The units and structures of writing, even though their materiality provides them with a certain degree of otherness and uniqueness in comparison with abstract linguistic units and acoustic units of speech, always reflect and relate to linguistic structures. If graphic marks do *not* relate to language, they are not writing. Yet, writing is also its own system. As such, however, like language, it is a structurally complex and compositional semiotic system: units from lower levels are combined to form larger units from hierarchically higher levels.

What, now, should and could a theory of writing achieve? Most importantly, it should provide generalized explanations of how writing systems work and *why* they work that way. Like arguably all theories, a theory of writing should strive to be based on empirical evidence, which, due to the lack of evidence in some areas, poses a challenge. In short, (even preliminary) evidence serves to draft the first version of a theory of writing, which can and needs to be revised when new relevant evidence is gathered.

The enterprise sketched thus far failed at the first step outlined above: the lack of a uniform method of describing diverse writing systems makes it impossible to compare research focusing on different writing systems and to correctly interpret, classify, and integrate new empirical evidence. Even within the disciplines participating in grapholinguistics, most notably linguistics, there is no consensus on the concepts and terminology pertaining to writing. We may not underestimate the effects this has also on grapholinguistic research carried out in other subjects such as psychology, neuroscience, etc., which rely on basic linguistic concepts. If not even linguists can agree on how to conceptualize writing, psychologists, neuroscientists, etc. cannot be expected to uniformly specify which structures of writing they are investigating in their studies. A shared vocabulary and theoretical framework are crucial. Otherwise, we may be speaking about the same phenomena (or different phenomena, for that matter) without being aware of it and sitting on undiscovered generalizations.

The modular model of writing systems developed by Neef (2015) is a promising step in establishing a descriptive basis for a theory of writing. Although it was not explicitly designed to be capable of describing the structural makeup of all writing systems, it can easily be generalized and modified to serve this purpose (cf. the introduction to Part II). Its subdivision in modules allows not only a distinction between parts of writing systems but also the associated logical separation of the respective grapholinguistic subbranches of graphetics, graphematics, and orthography. When treated in-depth, a description of these modules raises a central question to which there is not one definite and correct answer but a number of equally acceptable answers (cf. also Haspelmath 2010): how much unity is there in the diversity of writing systems? This question can also be asked the other way around: how much diversity is there in the unity of writing systems? Of these two questions, the first better captures what appears to have stalled theoretical progress in grapholinguistics: the perception that writing systems exhibit a degree of diversity that renders any attempt to strive for uniform explanations futile. This entails the opinion that any model that is able to account for diverse writing systems is too general to be of theoretical value.

However, at their core, all writing systems do function equally, circling back to what was posited above: they *all* represent language by graphic means. Structurally, too, there must be fundamental similarities: for instance, each writing system has a minimal unit. In the first step, the minimality of such a unit is evaluated graphetically (since writing is a visual/graphic modality and many of its systematics stem from this), independently of the linguistic unit that it may relate to. Here, empty spaces offer salient clues for a segmentation. In the next step, one can evaluate which linguistic units the identified smallest graphetic units correspond with. Trivially, then, we can identify a minimal graphematic unit in each writing system, which serves as the basis for the uniform concept of grapheme proposed in this book. The possibility of such a concept is not only rejected by those who claim there is altogether no grapheme but, ironically, also by those who aim to define the grapheme by analogy with other linguistic units such as the phoneme or the morpheme, as defining the grapheme analogously to a single of these units makes it inapplicable to writing systems that are not based on those units. In fact, linguistic diversity does not straightforwardly translate to written diversity. Vast differences in the structure of languages such as German and Chinese, for instance, do not necessarily imply that the differences are as drastic in the structural makeup of their respective writing systems.

In a comparative framework, the grapheme, as defined in this book (cf. Section 2.2), is the basic unit of writing. It a) differentiates meaning, b) relates to a linguistic unit, and c) is minimal. The type of linguistic unit that the default graphemes of a writing system correspond with is the so-called unit of representation (cf. Section 6.1), which has, quite uncontroversially, always been the basis of the assumption of different types in writing system typology (cf. Section 2.7).

By contrast, larger graphematic units are not as universal as the grapheme. For example, while some writing systems exhibit empty spaces that make visible word boundaries, others do not. The graphematic units that were postulated in the description of the German writing system – the graphematic syllable, the graphematic word, and the graphematic sentence – are thus examples of largely system-specific (or type-specific, i.e. alphabet-specific) units. The fact that they, although having been defined independently of linguistic units (phonological syllables, morphosyntactic words, and sentences of any kind), to an astonishing degree correspond with exactly those units is further evidence that writing represents language, which is the claim at the core of the narrow definition of writing. 'Deviances' – e.g. when a graphematic syllable does not correspond with a phonological syllable – remain, for the most part, exceptions. These exceptions do show, however, that beyond its function of representing language, writing remains a system of its own. It exhibits idiosyncratic traits that cannot be explained (solely) by its representational function but whose explanation rather lies within writing itself, highlighting the importance of a grapholinguistic methodology that is independent of the methods used in other linguistic subfields.

As for grapholinguistic diversity, the present study is characterized by several weaknesses: although it (at least partially) includes diachronic evidence of the development of writing systems as a crucial type of evidence, it is mainly focused on synchrony (cf. Elmentaler 2018 for a criticism of the synchronic nature of grapholinguistics). Thus, the writing systems that form the basis of this sketch of a functional theory of writing are modern writing systems that are still in use. Furthermore, they are writing systems that have been in continuous use for a long time. This is crucial insofar as it means they 'had the chance' to evolve naturally. This distinguishes them markedly from, on the one hand, ancient writing systems that are no longer in use and, on the other hand, writing systems that have been invented fairly recently and logically have not been in use for a long time. Both will need to be considered in further developments of a theory of writing, as their inclusion could either lead to a substantially different outcome and make necessary modifications or, potentially, serve as a reassurance that the theory outlined here is at least on the right track. Finally, many modern writing systems that fall into the same category as the ones featured prominently – e.g. German, Chinese, Arabic, Thai, Korean – are not yet studied as well and are thus characterized by a lack of available empirical data in many grapholinguistic subfields. This complicates or even makes impossible their inclusion in the theory until more data become available (see below).

A question that has not yet been explicitly addressed in this discussion is why there even is a need for a theory of writing. Indeed, we have done without one for such a long time that it appears as if there was indeed no pressing need for it. However, appearances are deceptive: it is not just Watt (1998) who, more than twenty years ago, lamented the lack of such a theory (as cited in the introduction of this book) but also many others. In fact, an abundance of experimental research in psychology, neuroscience, etc. also mentions the challenges that arise precisely due to a lack of theory and the lack of a straightforward methodology of grapholinguistic comparison. For instance, these challenges are often mentioned in the context of model-building, as models are hitherto most frequently based on alphabets (and specifically the English alphabet, cf. Share 2008, 2014). Another aspect that a theory could help elucidate is the investigation of isolated phenomena or features of writing: for instance, seeing the 'bigger picture' and thus knowing which parameters underlie the processing of writing systems and how they interact could help inform research on individual parameters such as word spacing. In other words, top-down knowledge of the larger context of how writing functions and how different features contribute could simultaneously help better understand the functions of its individual components.

### Meta-study on the state of grapholinguistic research

As a welcome side effect, this study showcased the breadth of research devoted to phenomena of writing and helped characterize grapholinguistics as the interdisciplinary field that subsumes all of this research (cf. Part I). Grapholinguistic research includes not only the various theoretical and descriptive approaches and traditions that stem mostly from linguistics or closely related fields but also the many types of empirical work that have been carried out and that have dealt with questions of writing. One of the challenges of collecting and reviewing such heterogeneous work dedicated to the subject of writing is that for a linguist (or a specialist in field x), it is much easier to evaluate linguistic research (or research of field x) on writing than it is to evaluate grapholinguistic research from different fields such as psychology or neuroscience. In their respective treatments of writing, these empirically driven fields have developed their distinct terms and theories for the subject, and the amount of psychological work on a multitude of facets of writing alone appears overwhelming. Accordingly, the research that ultimately contributed to the present sketch of a functional theory of writing is necessarily highly selective. The process of choosing it was both deductive and inductive and guided both by the fragments of the theory that were already established and the remaining gaps that these fragments laid bare. Naturalness Theory, as a well-developed (albeit nowadays marginal) linguistic theory, served as a crucial backdrop that provided a rough framework. The two strands of grapholinguistic research discussed in the following are theory and empiricism; I want to outline the state they are in and the work that must still be done. These two strands largely, but not completely, overlap with linguistics as the theoretical core of grapholinguistics and other participating disciplines – such as psychology, but also sociolinguistics, etc. – as empirical bases of grapholinguistics. This allocation is only rough, however, as linguistics has also offered empirical work; inversely, there exists, of course, an abundance of models and theory in empiricist psychology, just to name an example.

Regarding theory and description, modeling writing systems as consisting of subsystems or 'modules' is central: the first of these modules is the constitutive language system, all of whose levels (phonology, morphology, etc.) are the subject of different subfields of linguistics and are well-described. Since writing is defined as a modality of language, linguistic work represents the basis of every treatment of writing. The other three modules that are part of writing systems are the basis of a descriptive grapholinguistics and constitute their own respective subdisciplines. They are not as well-studied as language.

In fact, the first of them, graphetics, is heavily understudied and, in turn, underrepresented in grapholinguistic research (cf. Meletis 2015). On the one hand, this is not surprising, since, for a long time, questions concerning materiality were neglected by linguistics, which thus lags behind disciplines such as psychology that started much earlier to address questions about the effects the shape and appearance of writing have on processing. It is only in the last few decades and through the contact with neighboring (sub)disciplines such as media studies and the sociolinguistic treatment of typography that linguistics has opened the doors for questions of materiality. On the other hand, the lack of graphetic research and theory is startling given that the graphetic module is the most consolidated of all the modules and simultaneously the module that most clearly lends itself to a description by means of universal, script-independent categories (much like phonetics does for speech). The fact that graphetics is concerned solely with materiality rids it of any links to specific languages (of course, in a second step, the analysis of form-function correlations can be valuable). The overwhelming visual variety observable in the scripts of the world is thus deceitful as it conceals the fact that a universal graphetic theory is much more easily attained than a universal theory of graphematics or orthography and allows for system-independent generalizations. Consequently, many of the naturalness parameters described for the graphetic module are universal. This, ironically, makes the graphetic part of a functional theory of writing the best-established part from the get-go. However, while the present study laid a foundation, the description of the graphetic module, specifically with respect to a universal terminology of describing basic shapes in terms of features, is undeniably still in its infancy.

The second genuinely grapholinguistic module of writing systems, graphematics, has been treated mainly within the German grapholinguistic tradition; this results in an overrepresentation of works focusing on German graphematics and an inherent alphabetocentrism in the otherwise valuable establishment of graphematic concepts and terms. These concepts and terms, crucially, raise the above-mentioned question of universality vs. diversity. The comprehensive chapter on graphematics in this book (cf. Chapter 2) was thus not only concerned with critically summarizing previous graphematic work but also with overcoming the corset of language-specific categories by examining possible universals inherent in the graphematic module. Interestingly, following Daniels' (1991) rejection of graphematics, graphematic questions never really gained currency in the Anglo-American tradition of research on writing. This is also based on another problem, a problem of terminological and, more gravely, conceptual underdifferentiation: in English, when 'writing system' is meant, 'orthography' is often used instead. As argued in the course of this study, from a phenomenological point of view, orthography is indeed primary, which means in most cases, the data scholars of writing are dealing with are orthographic rather than graphematic. This, however, does not mean we should fail to acknowledge the crucial conceptual difference between the internal, systematic regularities of the graphematic module and the external, standardized, and (in most cases, English being an exception here) codified norms of the orthographic module.

This serves as the link to the last module of writing systems, orthography. It is also underrepresented in grapholinguistic research, also partially due to the above-mentioned fact that 'orthography' as a term has been appropriated descriptively. Given this situation, questions pertaining to orthography are – if treated at all – relegated to sociolinguistics, especially to the subfield of linguistic policy. Consequently, questions such as which properties of the graphetic and graphematic modules of typologically diverse writing systems lend themselves to orthographic regulation are seldom investigated. Suffice it to say that capitalization, for instance, is not as universal an orthographic concern as alphabetocentrism would have us believe. Neither is word separation.

Let us now turn to the empirical side of the equation. The transfer of the concepts of natural processes and naturalness parameters as well as the consideration of external evidence from Naturalness Theory (cf. Section 4.2) made available a rough framework that pointed to the grapholinguistic research necessary in developing a functional theory of writing. Accordingly, external evidence must come in the form of studies from a broad variety of grapholinguistic subdisciplines. Many works that help shape the theory stem, for instance, from psycholinguistics, psychology, and neuroscience, others from historical linguistics and philology, as well as sociolinguistics, just to list a few. While this breadth of research undoubtedly gives the theory its basis as, in using a range of diverse methods, it covers questions of children's literacy acquisition, on-line processing during reading and writing, the historical development of writing systems, literacy development in non-literate communities, and many others, there do remain several striking gaps with respect to evidence that is necessary to build a solid theory. These gaps point to general grapholinguistic desiderata that can be summarized as follows:

— *Research on a greater variety of writing systems.* Many writing systems are not well-studied. This includes their description and their inclusion in empirical research. In order to include a given writing system in a theory of writing, what is required is, aside from a description using comparative concepts such as the ones outlined in this book, evidence on (1) its historical development, (2) its acquisition by children, and (3) its processing both by unimpaired and impaired adults. Writing systems for which all these types of evidence are available are scarce. For many alphabets, such evidence is available. However, since this evidence then pertains to a sole type of writing system, its abundance is not necessarily an advantage for a theory of writing. Thus, evidence must be made available for a broad range of typologically distinct writing systems.

— *Research on the effect of naturalness parameters on processing.* For some of the naturalness parameters that were described in the course of this study, little to no empirical evidence exists on the effect they have (or do not have) on processing. For example, the effect of endophoric diagrammaticity, i.e. the systematicity of relationships between the units of a writing system, has not been empirically studied. While it is an obvious assumption that a highly diagrammatic writing system – such as Korean or Cree – is processed more efficiently than a writing system lacking such diagrammaticity (take, for example, Thai, cf. Section 6.2), such hypotheses can only be confirmed by evidence. The same goes for indexicality or compositional transparency. To reiterate what has been said above: to be able to separate universal from type-specific or even system-specific effects, research on these questions must be carried out for a variety of writing systems.

Given that several naturalness parameters were only systematically described for the first time in the context of this study, it is not surprising that there is no empirical evidence for them yet. The theoretical treatment of these parameters and their embedding in an overarching functional-theoretical framework would make visible their interaction with each other as well as their relevance for the description and processing of writing systems. This, ultimately, is one of the outcomes of the present study: highlighting gaps in research by attempting to develop the first sketch of a functional theory. Hope is that researchers from all grapholinguistic subfields will take note of these gaps and see them as incentives for further research.

What is an absolute necessity for a functional theory of writing to blossom and work in the long term is scholars' willingness to think outside their own boxes. This includes looking for evidence in places one would usually not look for it, and once such evidence is found, actually incorporating it in one's own research. It also means cooperating with scholars from vastly different disciplines and, finally, making available one's own findings to a broader audience than just one's own community. This is what I have attempted to do here, arguably with varying degrees of success. The utilization of evidence from all corners of the grapholinguistic sphere raises several obvious challenges: in the first step, as just mentioned, it is first necessary to even *find* it. As an example, take the studies of Garrod et al. (2007) and Caldwell & Smith (2012) that propose explanations of why pictography, despite its perceptual benefits, has decreased in writing systems and how sociocommunicative factors make this decrease possible. These studies had, to my knowledge, not yet been referred to in works on writing systems. On the one hand, this is likely because they are explanative rather than descriptive, and the grapholinguistic focus has thus far been on description. On the other hand, it is also because these studies are not obviously related to writing systems and a theory of writing. Thus, allowing to gain a better understanding of writing was not the goal of these studies but a side effect that can be utilized by scholars interested in writing. What these examples underline is that there are likely many more relevant findings 'out there', findings that I have missed, partially due to the fact that they were generated in fields that I did not even think of consulting. 'Undiscovered' research that can help flesh out a functional theory of writing must thus be found, operationalized, and integrated. This is where scholars from all fields who are interested in writing, all of them 'grapholinguists' to some degree, are invited to cooperate and to make available and promote their work more broadly. Interdisciplinary conferences on writing that make this possible are already a good start.

## A *tertium comparationis* for the comparison of writing systems

In the preceding chapters, several graphetic and graphematic categories were described, which, following terminology from Natural Morphology (cf. Section 4.2.3), were referred to as *parameters*. For graphetics, graphematics, and, to a lesser degree, orthography, parameters of systematic, linguistic, processing, and sociocultural nature were identified. Here, I want to discuss how they can inform the comparison of writing systems, what they mean for typology, and how naturalness conflicts and the interaction of the different levels of naturalness reveal a great deal about the nature of writing.

While the discussed parameters can be studied in isolation, the theoretical framework they are embedded in accounts for the fact that they interact with each other in complex ways. Accordingly, depending on the perspective adopted as well as the underlying epistemological interest, future grapholinguistic research can utilize these parameters in two different ways:

(1) In a *holistic approach*, the naturalness of entire (individual) writing systems can be studied and described. This amounts to in-depth analyses of writing systems in which they are elucidated from all possible perspectives. Here, the global question "why is writing the way it is?" is shifted to a local level and is reformulated to "why is this particular writing system the way it is?". The results of this endeavor are not traditional descriptions but rather descriptions that are explanatory given that they do not merely list features of writing systems but strive to explain the genesis of these features. In the next step, these elaborate accounts of individual writing systems can be compared, which is possible because their descriptions were made within a single theoretical framework. This, in turn, could potentially lead to a high degree of knowledge gain for grapholinguistics.

By contrast, in an (2) *atomistic approach*, individual parameters can be studied in detail. The graphetic parameter of *curvature* (or *roundness*), for example, can be studied in a way similar to how *cardinality* (cf. Morin 2018) or *topology* (cf. Changizi et al. 2006) have been studied, i.e. not just descriptively, amounting to a qualitative analysis as well as a quantitative analysis of this feature's distribution in the world's scripts, but also with respect to the effect the feature has on processing (cf. Section 7.1.3). Likewise, the graphematic parameter of *compositional transparency*, for instance, could benefit not only from a greater number of examples but more precise examples from a variety of typologically distinct writing systems. In an atomistic approach, thus, what is gained is a better understanding of individual parameters. This also automatically refines the theory and, in turn, the entire apparatus, i.e. the 'big picture' that is available for the above-mentioned holistic approach.

In both approaches, the parameters identified in this study can provide a starting point and a *tertium comparationis*. For instance, when two writing systems differ on one parameter, a rough first description of the naturalness configurations on said parameter as developed in this book can be used to evaluate which writing system is more natural with respect to that parameter. In this context, Roger's (1995) statement that some writing systems are 'better' than others is true, if only partially: it can be affirmed only for the local level, i.e. a given parameter. At the global level, by contrast, it appears unfeasible to evaluate whether an entire system is on the whole more natural than another system. It is, of course, possible that a system accumulates more naturalness on a greater number of parameters; this, at least quantitatively, gives the impression that this system is more natural as a whole. However, the question of whether all parameters are equivalent or whether, in a conflict between two parameters, possible outcomes are weighted differently, remains unanswered. In other words: is there a ranking of parameters (reminiscent of Optimality Theory)? This is one of the central questions for the future. Another possibility could be, of course, that all systems must be, if seen globally, equally natural as they are self-regulating systems and unnaturalness on some of their parameters is automatically compensated for by more naturalness on other parameters (cf. Fenk-Oczlon & Fenk 1995). In this view, systems only differ with respect to *where*, i.e. on which parameters they exhibit (un)naturalness. Holistically, however, they are all equally natural.

In some respects, grapholinguistic typology is still in its infancy. The findings of the different fits (systematic, processing, sociocultural) of the graphetic module provide interesting input for a prospective script typology (cf. also Section 1.3) but are inconclusive with respect to the question of which feature could serve as a base feature. One of the possible candidates that has emerged and does not require sophisticated descriptive work is quantitative visual complexity, a big part of which is so-called *perimetric complexity* (cf. Section 7.1.3). However, the question remains of whether classifying scripts as visually complex vs. less complex/ 'simple' scripts would result in a useful typology given that the complexity of basic shapes in a script interacts intricately with the size of a script, which in turn is in most cases determined by the type of a writing system, i.e. by graphematics. Complex scripts, thus, would likely turn out to be predominantly those scripts that are used for large syllabaries or scripts that are employed for morphographic writing systems. A different parameter that could potentially serve as a base parameter is *systematicity*, distinguishing systematic scripts from unsystematic scripts. Ultimately, however, a list of parameters for the comparison of scripts might suffice and there might be no apparent use for a typology based on a single one of them. Furthermore, as this book has revealed, the lack of descriptive work in graphetics is still remarkable, especially with respect to the description of basic shapes, as there is no uniform method for describing the basic shapes of the world's scripts yet. It is clear that such a description must proceed visually, and it is obvious that *spatiality* is of the utmost relevance, but given the visual variability of the tens of thousands of basic shapes, whether a unified qualitative method of description can be achieved remains doubtable (cf. Section 5.1). It might, however, not even be necessary to describe and compare basic shapes holistically. Several individual parameters cover different aspects such as the types of connections between elementary forms. Scripts could readily be compared with respect to those more fine-grained parameters. In any case, establishing a script typology and using it for further research is a cyclical process: new findings will influence existing typologies and these typologies (depending on whether they are received) will guide research.

'Writing system' typology, which more accurately should be termed 'graphematic typology', is much more advanced than a graphetic typology of scripts. However, it has been claimed that more fine-grained typological distinctions are necessary since the types of writing systems that have been assumed thus far fail to capture all of the ways in which writing systems can differ (cf. Weingarten 2011; Gnanadesikan 2017). This is where the graphematic naturalness parameters introduced in this study come into play. A promising parameter that has been used for minor refinements of phonographic types of writing systems in the past, particularly the alphabetic type, is graphematic *transparency*. For this parameter, what is most interesting is the relation between phonographic and morphographic transparency, as they conflict in languages with complex phonologies (i.e. a large amount of morphonology). In such cases, one type of transparency must necessarily be prioritized. Their interaction has led to the assumption of, for instance, a "morphophonographic type" in which morphography plays a crucial role but remains secondary to phonography (cf. DeFrancis 1989: 71; Hill 1967). In short, an assessment of the graphematic transparency (or transparencies) of different alphabets but also different abjads and abugidas can lead to a refinement of typological distinctions; the same can be achieved for non-segmental syllabaries and even morphographic writing systems, as long as phonography also assumes a certain role in them. The other graphematic naturalness parameters could be treated analogously. Thus, while I would not propose a wholly new typology of writing systems based on any one of the parameters of graphematic naturalness, they can at least modify existing typologies that are all based on the parameter of *unit of representation*. In any case, the parameters not only offer ways of refining existing typologies but can indeed also serve as base criteria for new typologies that are independent of the fundamental phonography/morphography distinction. With the help of *figure—ground*, for example, typological distinctions could describe how writing systems indicate word boundaries, with unspaced writing systems such as Chinese and Thai being located at one end of the spectrum and spaced writing systems such as alphabets at the other. Between them, on a continuum, there lie systems that do not exhibit spaces but are equipped with other features that mark word boundaries (such as script alternation in Japanese). Note that what must precede such a typologization based on any of the naturalness parameters is an evaluation of what the purpose of the resulting typology would be. It is only after this question has been answered reasonably that the parameters should be put to use. Then, however, the possibilities are manifold.

One of the central cornerstones of Naturalness Theory is the concept of naturalness conflict. This is true also for the functional theory of writing that draws on core concepts of Naturalness Theory. The most central conflict in writing is just a special variant of the struggle that is at the core of all communication: the conflict between sender and receiver, here between writer and reader. In this vein, various scholars claimed that with respect to the written modality, perception is primary. The evidence presented in this study supports this. A demonstrative example comes in the form of children's difficulties in producing modified links between two graphs in handwriting (cf. Gosse et al. 2018). 'Modified' pertains to the fact that when two basic shapes are produced in sequence in cursive handwriting, effects of coarticulation occur because the graphs that materialize the basic shapes are connected. On the basis of other evidence, continuity and connectedness of strokes and graphs had been identified as natural in the course of handwriting. The challenge posed to children by modified links, which assure a continuous production process, is thus not straightforwardly explainable. Not when perception is disregarded, that is. A possible, perception-based explanation is that basic shapes are stored as invariant visual representations, which would correspond with their descriptive definition as abstract visual configurations. Now, if the actual product, i.e. the graph that is produced, differs from the mental representation, this likely requires additional cognitive effort, since the visual invariance that is preferred by perception and cognition is violated. As Watt phrased it, "the program is merely a servant to the pattern" (cf. Watt 1988: 201) and a deviation from the pattern on the grounds of a modified program might cause the observed problems with modified links in children.

However, that is not the end of the story. Other evidence suggests that the hierarchy between perception and production is not fixed and that production can also exert a profound influence on perception. This was at least one of the possible explanations for the results of an experiment in which participants faced severe problems in retrieving the shape |g| from memory, producing it, and even recognizing it amongst a number of ill-formed distractors (cf. Wong et al. 2018). Unlike the alternative allograph |ɡ| that is also associated with the grapheme <g>, |g| is almost never produced in handwriting. The cognitive representation of the 'pattern', thus, even though it is visually ubiquitous in the graphetic environment that constantly surrounds users of Roman script, might be deprived of facilitative input from the 'program'.

Another vital finding of this sketch of a theory is that the different fits themselves are in conflict. As they are based on different types of linguistic or extralinguistic foundations, this points to more fundamental antagonisms at the heart not only of writing but also language. Strikingly, the fits also appear to strongly correlate with the different sublevels of naturalness. What is natural with respect to the linguistic fit depends mostly on the type of the language that is represented graphematically, meaning the linguistic fit correlates largely with typological naturalness. This is most obvious in the discussion about a fitting *unit of representation* and the fact that most of the other parameters are heavily influenced by the configuration of that parameter. By contrast, the systematic fit of scripts is assessed at the universal level: since the description of scripts and the evaluation of their systematicity is visual/graphic and independent of language, the descriptive principles and features remain the same regardless of the script in question. The graphetic and graphematic processing fits, too, are universal, i.e. system-independent in nature, since the physiological and cognitive makeup of humans does not differ with respect to the scripts and writing systems that are used. However, as evidenced by the fact that users of unspaced writing systems such as Chinese or Thai do not experience reading problems even though their writing systems exhibit unnaturalness with respect to the parameter of *figure—ground*, system-specific naturalness can be introduced or influenced by idiosyncratic features of individual systems, with users becoming familiar with these idiosyncrasies up to a point where they become natural for processing. Finally, the sociocultural fit is, unsurprisingly, most relevant at the system-dependent level. While parameters such as technological availability are universal, their values are system-specific. By contrast, the values of, for instance, the processing parameters are more universal. Thus, using leaves as a writing surface for the Burmese writing system might be natural system-dependently given that it is influenced by the material that is available. It is socioculturally determined features such as these that can override the descriptive fits (linguistic and systematic) as well as the processing fit and decrease their naturalness. In fact, the parameters of the sociocultural fit lead to a situation in which the questionable and system-oriented claim "every language gets the writing system it deserves" could be rephrased as the more user-oriented claim "every literate community gets the writing system it deserves"; this thought will be revisited in the conclusion below.

# V Conclusion

> I regard a natural graphematics [in the broad sense of 'natural grapholinguistics', D.M.] not only a possible, but a promising way of deepening the description of phenomena of writing and reaching explanations about their development. However, for this purpose, the empirical basis is – in comparison to classical fields of linguistics – still very meager.
>
> (Munske 1994: 22, my translation)[275]

It is high time that writing is given the elaborate scholarly treatment that it deserves. The present study is a step in that direction. It is fragmentary, and it raises more questions than it can answer, but it is arguably nonetheless promising as it highlights how a theory that was initially designed for language can be productively extended and modified to investigate and explain the nature of writing, which is not only a modality of language but also a system of its own complete with idiosyncratic features. Accordingly, the main goal of the present enterprise was to lay out clearly the steps that must be taken to gain a better understanding of the nature of writing.

In the introduction of this book, Watt's (1998: 118) quest for answers to questions such as "why each such writing system is the way it is, instead of some other way, and why all such systems have in common what they have in common" was quoted. These questions constitute the very heart of grapholinguistics. While answers to these questions are necessarily multilayered and elaborate, their gist is straightforward, almost trivially so: human physiology, cognition, and communication have shaped and continue to shape writing. These three cornerstones not only make each writing system what it is but are also responsible for the common core shared by the world's diverse writing systems. The same is true for language in general, including the thousands of languages of the world. Accordingly, the "central goal of linguistic theory is to shed light on the core of grammatical principles that is common to all languages", as Kager (1999: 1) declares, adding that "a broad picture emerges of 'unity and

275 „[…] daß ich eine natürliche Graphematik nicht nur für möglich, sondern für einen vielversprechenden Weg halte, die Beschreibung von Phänomenen der Schriftlichkeit zu vertiefen und zum Erklären ihrer Entwicklungsprozesse vorzudringen. Allerdings ist hierfür die empirische Grundlage – im Vergleich zu den klassischen Bereichen der Systemlinguistik – noch sehr schmal."

variety'" from a large body of linguistic research. Given its comparative nature, the proposed functional theory of writing has the potential to reach precisely that goal, and indeed, as several findings discussed throughout this study have underlined, the common core of writing can be uncovered if we have both a fitting theoretical framework and the necessary empirical evidence. And, as a side note: when compared with the search for universals in the world's thousands of languages, the investigation of universals of and in writing proves much more manageable.

On the one hand, this is because writing is a much younger phenomenon than language. The visual richness of the world's scripts might be off-putting, giving the impression that writing systems are as diverse as the languages and writing systems they provide with a visual manifestation; this assumption, however, is a fallacy. Behind the visual variety lie a central common core and a guiding thread that can be identified: writing systems need to fit their respective languages, and more importantly, they need to fit their users, specifically their (processing) needs and (communicative) wishes. Although different scripts and writing systems have developed in varying directions over time, their core remains the same, as do the key factors in the history of writing. When it comes to these three subcomponents of grapholinguistic naturalness, when integrated into a theoretical framework, the linguistic fit constitutes a predominantly typological matter, while the processing fit is system-independent as it pertains to physiology and cognition, and the sociocultural fit, unsurprisingly, functions at the system-dependent level. In all these respects, writing echoes language. This is also not a groundbreaking finding given that writing, as a semiotic system based on language, is logically inseparably tied to language. This means that many of the steps that are necessary for the establishment of theories of language are also required in carving out a theory of writing. That way, the study of writing is like the study of language – but in a microcosm. It is because of these obvious parallels between language and writing that one does not need to start from scratch in the establishment of a theoretical framework for writing. As outlined in this book, linguistic Naturalness Theory appears to be a suitable match for discovering explanations for the nature of writing. However, it bears repeating that the specific linguistic (or possibly other, e.g. psychological) theory used as a starting point for a theory of writing is not fixed, and other approaches might provide equally intriguing bases for a treatment of writing. A combination or comparison of the suitability of these approaches for the subject of writing is definitely a fruitful endeavor for the future. In any case, and this is a fact, writing must be treated theoretically in some way and grapholinguistics should strive for explanation. When the pieces of the puzzle that is the scattered knowledge of writing are put together, the result can be a very sound picture. The sounder, the better the theory.

What follows is a fictional scenario that illustrates how the proposed functional theory of writing could provide explanations. Imagine a person (as a placeholder for an entire community) who wants to invent a writing system for her language that allows written communication between the members of her community. The first question she is confronted with is whether she adopts an existing script or invents a new script instead. As was argued in this book, this decision is influenced by several factors: does the person have prior knowledge of what writing is, i.e. that it already exists elsewhere, and is used for other languages? Is the person possibly already literate in a different writing system? Or is someone who is literate from 'outside' helping her create the new writing system? Additional factors pertain to the availability of resources: is technology (interpreted very broadly) available in the person's region? In our modern times, are there computers, smartphones, is there a connection to the internet? If existing scripts (such as Roman script) are familiar to the linguistic community in question and thus represent possible choices but are associated with different cultures that are, for some reason, negatively connoted, this might prove to be a sociocultural knock-out criterion. Finally, for various reasons (including the above-mentioned ones), the person decides to invent an entirely new script. In this special scenario of the creation of a script custom-tailored for the language the person intends to 'write down', a decision must be made on a unit of representation, i.e. the linguistic units that the basic units of the new writing system – the graphemes – shall relate to. Making this decision requires sophisticated metalinguistic work that the person might perform consciously or unconsciously. It includes questions such as: what is the language's structure? What level of representation is most feasible? In the case of an analytic morphosyllabic language, the person might decide on a morphographic system, also since she, as she is not literate in any writing system, likely exhibits no phoneme awareness, making phonological segments (for the most part) inaccessible to her. The drawbacks of large inventories of basic shapes and graphemes that are caused by morphography are outweighed by the possibility of pictography, the script creator realizes, which will allow her fellow community members to learn and memorize the system's units quickly. The shapes she creates are pictographic, at least for words and morphemes that refer to concrete objects. By contrast, for abstract concepts, she must come up with arbitrary shapes. While they have perceptual benefits, producing pictographic basic shapes is cumbersome, as it resembles drawing, and this means the writing process is not yet as efficient and quick as it could be. Furthermore, because it poses a challenge to think of as many distinct basic shapes as there are morphemes in the language, the script inventor instead reuses some of the existing shapes by inverting them horizontally, which introduces extrinsic symmetry into the system. This will help her remember how shapes are produced, and she effectively needs to memorize fewer shapes. The shapes themselves are rather curved in nature since what is available as material in the region is palm leaves and spiky stems from other plants that are used to carve on the leaves. When she finally unveils the writing system to her community, the other members start to learn and use it. During the acquisition process, they make mistakes on the extrinsically symmetrical shapes that have different graphematic values, reversing them so that they are oriented in the direction in which most other shapes in the script are oriented, which is also the direction of writing and reading. When later, paper and pens become available as writing materials, the production process becomes quicker, and shapes become connected and more cursive. In this process, they gradually lose their pictographic character, which is not a problem for communication, however, since the complex communicative interaction between the now fully literate members of the community establishes conventions: from a semiotic point of view, symbols supersede icons. When at a later point, the language in question comes into contact with a different language and words borrowed from this foreign language must be written in the native writing system, users start using the originally morphographic graphemes only for their phonological value, i.e. their pronunciation. At another point in time, the script may be encoded in Unicode. It can then be used digitally, and typefaces are created; keyboards and other technology become available and basic shapes are cemented in their appearance and become visually stable. Due to the above-mentioned mixing of morphography and phonography caused by the integration of foreign material, a large degree of graphematic variation is introduced into the system, i.e. the possibilities of spelling given utterances are many. Thus, a committee is formed to decide on an orthographic standard. Over the course of time, the spoken modality of the language develops more quickly than the written one, and the orthography needs to be reformed to react to this discrepancy. This story, of course, continues for as long as the writing system is in use.

All developments in this scenario can be explained in the functional theoretical framework proposed in this book. It accounts for the development and the predictable parts of the change of scripts and writing systems, the frequency of phenomena, errors that are made, factors that influence sociocultural preferences, etc. It is not perfect, and in many cases, it will run into problems and lack explanations for certain (unpredictable) phenomena, and this is precisely where grapholinguistic theorizing needs to continue.

Note that the scenario above represents, synchronically, the most marginal case. Nowadays, unsophisticated grammatogenies, i.e. creations of writing systems by people who do not know that writing already exists and/or are not literate in an existing system, are rare. Scripts are extremely seldom invented and are nowadays commonly not written on leaves, since in most cases, other material is available. More pressing and equally explainable scenarios for a naturalist treatment of writing are those in which scripts and graphematic principles are adopted and adapted from one system to another, leading to complex interactions between donor and target systems and making imperative a consideration of traces that the donor system leaves in the target system. The distinction between systems invented from scratch and systems that were heavily influenced or shaped by existing systems is central and leads to crucial differences in how naturalness is reflected in a writing system (cf. also Sampson 2016). This observation is paramount in going forward with a functional theory of writing.

Two provocative and persistent questions about writing recurred throughout this book. The first one was whether some writing systems are better than others (cf. Rogers 1995: 31). A very simplistic answer to this is: *yes*, some systems appear to be better than others. However, the answer is never absolute, as an investigation of this question in an absolute manner is unfeasible. This means that it is not entire systems that can be reasonably compared but instead the question of how these systems cope with various problems, including how systematic they are (*systematic fit*), how they represent their respective underlying language (*linguistic fit*), how they are visually, graphomotorically, and cognitively suited for our human capacities (*processing fits*), and how they socially, culturally, ideologically (and politically, religiously, etc.) index and represent what the communities who use them want them to convey (*sociocultural fits*). An isolated claim such as "morphographic systems are inferior to phonographic ones" (Jones & Mooney 2017: 13) is, simply put, meaningless and, thus, false when no context is given. This question can only be investigated with a complex network of factors in mind and is only reasonable in relation to a specific situation or scenario: in case x, morphography might be less suited, but in case y, it might be more natural than phonography for a variety of factors. No entire system (or type of system) is better than another system in every respect. In fact, due to naturalness conflicts, a completely natural writing system cannot exist. These conflicts are dealt with in various ways across different systems.

The second of the above-mentioned questions is intricately related to the linguistic fit: does every language get the writing system it deserves (cf. Frost 2012: 266)? In short: *no*. If the underlying language were the most crucial and not just one of several conflicting variables, a writing system with a very natural linguistic fit (again, only on a number of parameters, never all, which is impossible) would be imaginable. However, what is most natural systematically and linguistically is not necessarily natural for processing, and – which is the crucial point in the answer to this question – for the sociocultural environment and a given community's needs. It is writing's communicative function and embedding in larger contexts that are not only linguistic but social and that commonly hinder languages from getting precisely the writing system that they would, from a strictly linguistic perspective, 'deserve'. However, if the four fits align and sociocultural factors reinforce what is natural linguistically, which, for instance, is arguably the case for the writing system of Korean, the impression can arise that languages can indeed get what they deserve. But, more importantly, so do the literate communities in these cases. Since sociocultural factors, which include orthographic regulation and standardization, so often 'have the last say', the question cited above should be reformulated as 'does every literate community get the writing system it deserves?'. Given the dominant nature of the sociocultural fit, the answer to this modified question is more likely to be affirmative since the wishes of the community (or political authority, etc.) are shaping the system to begin with.

# VI Continuation

In the discussion (cf. Part IV above), several points were mentioned that need to be elaborated further in the context of future research. Most importantly, although many writing systems have already been described, they should be 're-described' within a uniform theoretical framework or at least integrated into such a framework that affords shared concepts and terminology, as this is how comparability can be assured. Indeed, and this bears repeating, comparison is the crux of a theory of writing. This extends also to ancient writing systems that are no longer in use as well as modern writing systems that have been invented recently, both of which were underrepresented in the present study. Consequently, as the body of descriptive grapholinguistic work gradually grows and writing systems that have altogether not yet been described are treated for the first time, not only an explanatory functional theory must be refined but also the descriptive foundation at its core. Concepts such as *basic shape* or *grapheme*, as final as their definition might appear at this point (given their level of generality), are by no means set in stone (to use a writing-related metaphor), and if (or when) writing systems are encountered that cannot be accommodated by them, they will need to be modified and redefined. Grapholinguistics is young, and it is in a state of flux.

There are a number of tasks to be approached in the near future. Firstly, the research that is already 'out there' and scattered across disciplines as well as academic cultures around the world and that is published in a variety of different languages must be located and integrated into the theory. As this study has illustrated, much of the existing research was implicitly naturalist (or more generally: 'functional') in its methodology and assumptions and, crucially, its striving for explanations. The extensive treatment of the complex phenomenon of writing within the scope of a PhD thesis (which this book is based on) expectedly resulted in a fragmentary picture, which means much more work is required to find existing evidence from the heterogeneous fields that are part of the interdisciplinary grapholinguistics. Some types of evidence such as evidence from disturbances of reading and writing must be consulted more elaborately, and some fields such as literacy instruction must be included to begin with, to name only two examples of what has not yet been done. However, a first outline of the global picture of why writing is the way it is was needed for a start, and going forward, it will at least prove useful in distinguishing the research that is relevant for a theory from research that is not.

Aside from existing research and evidence, more empirical evidence is required for most scripts and writing systems with respect to their systematic, linguistic, processing, and sociocultural fits. For example, it is astonishing that there does not appear to exist – at least not in English – any research on the processing of the Cree or Georgian writing systems. Such evidence could provide explanations for graphetic parameters such as *distinctiveness* (or the lack of distinctiveness due to *extrinsic symmetry*) and, in the case of Georgian, *curvature*, given that the basic shapes of Georgian are predominantly curved.

In addition to more research, intensified contact between different fields is necessary, and findings must be made available in a global framework. For a theory of writing to advance, everyone interested in written phenomena must be able to become involved, and this necessitates a common ground, i.e. a presentation of findings on writing that is more accessible to all interested and invested scholars of the various grapholinguistic subfields. Finally, case studies of writing systems are necessary that test the framework as presented in this study. In the present context, examples of parameters were given from a multitude of systems, but no single system was evaluated with respect to all parameters. However, such a comprehensive evaluation of an entire system would be an important test for the theory as it likely highlights gaps in the theory or where it requires revision. Such case studies of the naturalness of different scripts and/or writing systems that are carried out in the same theoretical framework could, in a further step, also be straightforwardly compared, which would allow attaining a deeper understanding of universal, typological, and system-specific aspects of writing.

This outline of a grapholinguistic theory offers a collection of ideas about writing, a status report about relevant research, a discovery of desiderata, and a new perspective. It is a start, but most importantly, it is an invitation.

# VII References


Abdelhay, Ashraf, Busi Makoni (neé Dube) & Sinfree Makoni (2018): When God is a linguist: missionary orthographies as a site of social differentiation and the technology of location. In Constanze Weth & Kasper Juffermans (eds.), *The tyranny of writing: ideologies of the written word*, 97–112. London, New York: Bloomsbury.

Abu-Rabia, Salim (1998): Reading Arabic texts: effects of text type, reader type and vowelization. *Reading and Writing* 10.2: 105–119. DOI: 10.1023/A:1007906222227

Abu-Rabia, Salim (2001): The role of vowels in reading Semitic scripts: data from Arabic and Hebrew. *Reading and Writing* 14.1–2: 39–59. DOI: 10.1023/A:1008147606320

Adam, Isabell (2013): Graph. In Martin Neef, Said Sahel & Rüdiger Weingarten (eds.), *Schriftlinguistik/Grapholinguistics* (= Dictionaries of Linguistics and Communication Science, 5). Berlin, Boston: De Gruyter. DOI: 10.1515/wsk.5.0.graph

Ágel, Vilmos (2003): Prinzipien der Grammatik. In Anja Lobenstein-Reichmann & Oskar Reichmann (eds.), *Neue historische Grammatiken. Zum Stand der Grammatikschreibung historischer Sprachstufen des Deutschen und anderer Sprachen* (= Reihe Germanistische Linguistik, 243), 1–46. Berlin, Boston: De Gruyter. DOI: 10.1515/9783110913194.1

Albrow, Kenneth H. (1972): *The English writing system: Notes towards a description*. London: Longman.

Alessandrini, Jean (1979): Nouvelle classification typographique: Codex 1980. *Communication et Langages* 43: 35–56.

Almog, Guy (2018): Getting out of hand? Examining the discourse of 'character amnesia'. *Modern Asian Studies*. DOI: 10.1017/S0026749X1700035X

Althaus, Hans P. (1973): Graphetik. In Hans P. Althaus, Helmut Henne & Herbert E. Wiegand (eds.), *Lexikon der germanistischen Linguistik*, 105–118. Tübingen: Niemeyer.

Altmann, Gabriel (2004): Script complexity. *Glottometrics* 8: 64–74.

Altmann, Gerry T. M., Alan Garnham & Yvette Dennis (1992): Avoiding the garden path: eye movements in context. *Journal of Memory and Language* 31.5: 685–712. DOI: 10.1016/0749-596X(92)90035-V

Amtliche Regelung (2018): *Regeln und Wörterverzeichnis. Aktualisierte Fassung des amtlichen Regelwerks entsprechend den Empfehlungen des Rats für deutsche Rechtschreibung 2016.* Redigierte Fassung vom Januar 2018. Mannheim: Rat für deutsche Rechtschreibung. <http://www.rechtschreibrat.com/DOX/rfdr_Regeln_2016_redigiert_2018.pdf> (September 22nd, 2020).

Anbar, Ada (1986): Reading acquisition of preschool children without systematic instruction. *Early Childhood Research Quarterly* 1.1: 69–83. DOI: 10.1016/0885-2006(86)90007-4

Anderson, Richard C., Wenling Li, Yu-Min Ku, Hua Shu & Ningning Wu (2003): Use of partial information in learning to read Chinese characters. *Journal of Educational Psychology* 95.1: 52–57. DOI: 10.1037/0022-0663.95.1.52

Androutsopoulos, Jannis (2000): Non-standard spellings in media texts: the case of German fanzines. *Journal of Sociolinguistics* 4.4: 514–533. DOI: 10.1111/1467-9481.00128

Aronoff, Mark (1992): Segmentalism in linguistics: the alphabetic basis of phonological theory. In Pamela Downing, Susan D. Lima & Michael Noonan (eds.), *The linguistics of literacy* (= Typological Studies in Language, 21), 71–82. Amsterdam: John Benjamins. DOI: 10.1075/tsl.21.09aro

Asadi, Ibrahim A. (2017): Reading Arabic with the diacritics for short vowels: vowelised but not necessarily easy to read. *Writing Systems Research* 9.2: 137–147. DOI: 10.1080/17586801.2017.1400493

Assmann, Jan (1988): Im Schatten junger Medienblüte. Ägypten und die Materialität des Zeichens. In Hans Gumbrecht & K. Ludwig Pfeiffer (eds.), *Materialität der Kommunikation*, 141–160. Frankfurt a. M.: Suhrkamp.

Augst, Gerhard (ed.) (1986): *New trends in graphemics and orthography*. Berlin, Boston: De Gruyter. DOI: 10.1515/9783110867329

Bahadori, Abdulfazi (1993): Alphabet in the boiling pot of politics. *Azerbaijan International* 1.3: 10–13. <https://www.azer.com/aiweb/categories/magazine/13_folder/13_articles/13_alphapolitics.html> (January 23rd, 2019).

Bahlo, Nils, Tabea Becker & Daniel Steckbauer (2016): Von „Klugscheißern" und „Grammatik-Nazis" – Grammatische Normierung im Internet. In Carmen Spiegel & Daniel Gysin (eds.), *Jugendsprache in Schule, Medien und Alltag*, 275–286. Frankfurt a. M.: Peter Lang.

Bai, Xuejun, Guoli Yan, Simon P. Liversedge, Chuanli Zang & Keith Rayner (2008): Reading spaced and unspaced Chinese text: evidence from eye movements. *Journal of Experimental Psychology* 34.5: 1277–1287. DOI: 10.1037/0096-1523.34.5.1277

Bailey, Charles-James N. (1984): The concept of balance and linguistic naturalness. *Language Sciences* 6.2: 229–238. DOI: 10.1016/S0388-0001(84)80018-2

Balestra, Miriam B. (2017): *Vokalschreibungen im Deutschen: Eine graphotaktische Analyse einsilbiger Nomen*. Frankfurt a. M.: Peter Lang. DOI: 10.3726/b11532

Balestra, Miriam B., Annalen Appelt & Martin Neef (2014): Systematische Beschränkungen für Schreibungen im grammatischen Wortschatz des Deutschen: der Konsonant [f]. *Zeitschrift für Sprachwissenschaft* 33.2: 129–163. DOI: 10.1515/zfs-2014-0006

Bar, Moshe & Maital Neta (2006): Humans prefer curved visual objects. *Psychological Science* 17.8: 645–648. DOI: 10.1111/j.1467-9280.2006.01759.x

Baroni, Antonio (2011): Alphabetic vs. non-alphabetic writing: linguistic fit and natural tendencies. *Rivista di Linguistica* 23.2: 127–159.

Baroni, Antonio (2016): Constraint interaction and writing systems typology. *Dossiers d'HEL, SHESL* 9: 290–303.

Barton, David (1995): Some problems with an evolutionary view of written language. In Stanisław Puppel (ed.), *The biology of language*, 19–30. Amsterdam, Philadelphia: John Benjamins. DOI: 10.1075/z.75.04bar

Barton, David (2007): *Literacy: an introduction to the ecology of written language*, 2nd edition. Oxford: Wiley-Blackwell.

Baudusch, Renate (1976): Einige Bemerkungen zur geltenden Regelung unserer Zeichensetzung. *Sprachpflege* 10: 196–201.

Bauernschmidt, Amy (1980): The ideal orthography. *Notes on Literacy* 32: 12–21.

Beard, Robert (1994): Morphology: History. In R. E. Asher & J. M. Y. Simpson (eds.), *The encyclopedia of language and linguistics*, 2573–2576. Oxford: Pergamon Press.

Behr, Wolfgang (2010): In the interstices of representation: Ludic writing and the locus of polysemy in the Chinese sign. In Alex de Voogt & Irving L. Finkel (eds.), *The idea of writing: play and complexity* (= The Idea of Writing, 1), 281–314. Leiden: Brill. DOI: 10.1163/ej.9789004174467.i-396.94

Behr, Wolfgang (2018): 'Monosyllabism' and some other perennial clichés about the nature, origins and contacts of the Chinese language in Europe. In Angelika Malinar & Simone Müller (eds.), *Asia and Europe – interconnected: agents, concepts, and things*, 155–209. Wiesbaden: Harrassowitz.

Behrens, Ulrike (1989): *Wenn nicht alle Zeichen trügen. Interpunktion als Markierung syntaktischer Konstruktionen* (= Arbeiten zur Sprachanalyse, 9). Frankfurt a. M.: Peter Lang.

Berg, Kristian (2012): Identifying graphematic units: vowel and consonant letters. *Written Language and Literacy* 15.1: 26–45. DOI: 10.1075/wll.15.1.02ber

Berg, Kristian (2016a): Graphematische Variation. In Birgit Mesch & Christina Noack (eds.), *System, Norm und Gebrauch – drei Seiten derselben Medaille? Orthographische Kompetenz und Performanz zwischen System, Norm und Empirie*, 9–23. Baltmannsweiler: Schneider Hohengehren.

Berg, Kristian (2016b): Graphemic analysis and the spoken language bias. *Frontiers in Psychology* 7: 388. DOI: 10.3389/fpsyg.2016.00388

Berg, Kristian (2019): *Die Graphematik der Morpheme im Deutschen und Englischen*. Berlin, Boston: De Gruyter. DOI: 10.1515/9783110604856

Berg, Kristian, Beatrice Primus & Lutz Wagner (2016): Buchstabenmerkmal, Buchstabe, Graphem. In Ulrike Domahs & Beatrice Primus (eds.), *Handbuch Laut, Gebärde, Buchstabe* (= Handbooks of Linguistic Knowledge, 2), 337–355. Berlin, Boston: De Gruyter. DOI: 10.1515/9783110295993-019

Berg, Kristian & Mark Aronoff (2017): Self-organization in the spelling of English suffixes: The emergence of culture out of anarchy. *Language* 93.1: 37–64. DOI: 10.1353/lan.2017.0000

Berg, Kristian & Martin Evertz (2018): Graphematik – die Beziehung zwischen Sprache und Schrift. In Stefanie Dipper, Ralf Klabunde & Wiltrud Mihatsch (eds.), *Linguistik – Eine Einführung (nicht nur) für Germanisten, Romanisten und Anglisten*, 187–195. Berlin: Springer. DOI: 10.1007/978-3-662-55589-7_10

Berkemeier, Anne (1997): *Kognitive Prozesse beim Zweitschrifterwerb* (= Arbeiten zur Sprachanalyse, 29). Frankfurt a. M.: Peter Lang.

Berry, Jack (1958): The making of alphabets. In Eva Sivertsen (ed.), *Proceedings of the 8th International Congress of Linguists*, 752–764. Oslo: Oslo University Press.

Bertacca, Antonio (2002): Description and explanation of language change. In Katarzyna Dziubalska-Kołaczyk & Jarosław Weckwerth (eds.), *Future challenges for Natural Linguistics* (= LINCOM Studies in Theoretical Linguistics, 30), 1–23. München: LINCOM.

Bhatt, Parth M. (1988): Graphic systems, phonic systems, and linguistic representations. In Derrick de Kerckhove & Charles J. Lumsden (eds.), *The alphabet and the brain: the lateralization of writing*, 106–120. Berlin, Heidelberg: Springer. DOI: 10.1007/978-3-662-01093-8 _8

Bhide, Adeetee (2015): Early literacy experiences constrain L1 and L2 reading procedures. *Frontiers in Psychology* 6: 1446. DOI: 10.3389/fpsyg.2015.01446

Bhuvaneshwari, B. & Prakash Padakannaya (2014): Reading in Tamil: a more alphabetic and less syllabic akshara-based orthography. In Heather Winskel & Prakash Padakannaya (eds.), *South and Southeast Asian psycholinguistics*, 192–201. Cambridge: CUP. DOI: 10.1017/CBO9781139084642.022

Biber, Douglas (1988): *Variation across speech and writing*. Cambridge: CUP.

Birk, Elisabeth (2013): Graphem. In Martin Neef, Said Sahel & Rüdiger Weingarten (eds.), *Schriftlinguistik/Grapholinguistics* (= Dictionaries of Linguistics and Communication Science, 5). Berlin, Boston: De Gruyter. DOI: 10.1515/wsk.5.0.graphem

Blesser, Barry A., Robert J. Shillman, Theodore T. Kuklinski, Charles Cox, Murray Eden & J. Ventura (1974): A theoretical approach for character recognition based on phenomenological attributes. *International Journal of Man-Machine Studies* 6.6: 701–714. DOI: 10.1016/S0020-7373(74)80033-7

Bohn, Hartmut (1998): *Quantitative Untersuchungen der modernen chinesischen Sprache und Schrift* (= PHILOLOGIA – Sprachwissenschaftliche Forschungsergebnisse, 24). Hamburg: Dr. Kovač.

Boltz, William G. (1986): Early Chinese writing. *World Archaeology* 17.3: 420–436. DOI: 10.1080/00438243.1986.9979980

Boltz, William G. (2006): Pictographic myths. *Bochumer Jahrbuch zur Ostasienforschung* 30: 39–54.

Bonneton-Botté, Nathalie, Florence Bara, Nathalie Marec-Breton, Fanny De La Haye-Nicolas & Corentin Gonthier (2018): Perception of the cursive handwriting movement in writers and pre-writers. *Reading and Writing* 31.4: 927–943. DOI: 10.1007/s11145-018-9819-8

Boone, Elizabeth H. (1994): Introduction: writing and recording knowledge. In E. H. Boone & W. D. Mignolo (eds.), *Writing without words: alternative literacies in Mesoamerica and the Andes*, 3–26. Durham: Duke University Press.

Boone, Elizabeth H. (2004): Beyond writing. In Stephen D. Houston (ed.), *The first writing: script invention as history and process*, 313–348. Cambridge: CUP.

Booth, Robert W. & Ulrich W. Weger (2013): The function of regressions in reading: backward eye movements allow rereading. *Memory & Cognition* 41.1: 82–97. DOI: 10.3758/s13421-012-0244-y

Borleffs, Elisabeth, Ben A. M. Maassen, Heikki Lyytinen & Frans Zwarts (2017): Measuring orthographic transparency and morphological-syllabic complexity in alphabetic orthographies: a narrative review. *Reading and Writing* 30.8: 1617–1638. DOI: 10.1007/s11145-017-9741-5

Bosman, Anna M. T. & Guy C. Van Orden (1997): Why spelling is more difficult than reading. In Charles E. Perfetti, Laurence Rieben & Michel Fayol (eds.), *Learning to spell: research, theory, and practice across languages*, 173–194. Mahwah: Lawrence Erlbaum.

Bredel, Ursula (2008): *Die Interpunktion des Deutschen: Ein kompositionelles System zur Online-Steuerung des Lesens* (= Linguistische Arbeiten, 522). Tübingen: Niemeyer. DOI: 10.1515/9783484970502

Bredel, Ursula (2009): Das Interpunktionssystem des Deutschen. In Angelika Linke & Helmuth Feilke (eds.), *Oberfläche und Performanz* (= Reihe Germanistische Linguistik, 283), 117–136. Berlin, Boston: De Gruyter. DOI: 10.1515/9783484971240.2.117

Bredel, Ursula (2011): *Interpunktion* (= Kurze Einführungen in die germanistische Linguistik, 11). Heidelberg: Winter.

Bredel, Ursula (2015): Leerzeichen. In Martin Neef, Said Sahel & Rüdiger Weingarten (eds.), *Schriftlinguistik/Grapholinguistics* (= Dictionaries of Linguistics and Communication Science, 5). Berlin, Boston: De Gruyter. DOI: 10.1515/wsk.5.0.leerzeichen

Bredel, Ursula, Christina Noack & Ingo Plag (2013): Morphologie lesen. Stammkonstanzschreibung und Leseverstehen bei starken und schwachen Lesern. In Martin Neef & Carmen Scherer (eds.), *Die Schnittstelle von Morphologie und geschriebener Sprache* (= Linguistische Arbeiten, 551), 211–249. Berlin, Boston: De Gruyter. DOI: 10.1515/9783110336665.211

Brekle, Herbert E. (1994a): *Die Antiqualinie von ca. –1500 bis ca. +1500. Untersuchungen zur Morphogenese des westlichen Alphabets auf kognitivistischer Basis*. Münster: Nodus Publikationen.

Brekle, Herbert E. (1994b): Die Buchstabenformen westlicher Alphabetschriften in ihrer historischen Entwicklung. In Hartmut Günther & Otto Ludwig (eds.), *Schrift und Schrift-*

*lichkeit/Writing and its use* (= Handbooks of Linguistics and Communication Science, 10.1), 171–204. Berlin, Boston: De Gruyter. DOI: 10.1515/9783110111293.1.2.171

Brekle, Herbert E. (1994c): Some thoughts on a historico-genetic theory of the lettershapes of our alphabet. In W. C. Watt (ed.), *Writing systems and cognition* (= Neuropsychology and Cognition, 6), 129–139. Dordrecht: Springer. DOI: 10.1007/978-94-015-8285-8_8

Brekle, Herbert E. (1995): Neues über Groß- und Kleinbuchstaben. Theoretische Begründung der Entwicklung der römischen Majuskelformen zur Minuskelschrift. *Linguistische Berichte* 155: 3–21.

Brekle, Herbert E. (1996): Warum schreiben wir rechtsläufig? oder: Schreibfunktionale Zusammenhänge zwischen der Vektorialisierung von Buchstabenformen, der buchstabeninternen Produktionsrichtung (sinistral/dextral) und Schreibrichtung (sinistro-/dextrograd) in phönizischer und altgriechischer Zeit. *Linguistische Berichte* 166: 483–491.

Brekle, Herbert E. (1998): Randbedingungen und Gesetzmäßigkeiten im historischen Entwicklungsprozeß unserer Buchstabenformen. *OBST* 56: 1–10.

Brem, Silvia & Urs Maurer (2016): Lesen als neurobiologischer Prozess. In Ursula Rautenberg & Ute Schneider (eds.), *Lesen: ein interdisziplinäres Handbuch*, 117–140. Berlin, Boston: De Gruyter. DOI: 10.1515/9783110275537-008

Brody, Michal (2004): *The fixed word, the moving tongue: variation in written Yucatec Maya and the meandering evolution toward unified norms.* PhD dissertation, The University of Texas at Austin.

Brosh, Hezi (2015): Arabic spelling: errors, perceptions, and strategies. *Foreign Language Annals* 48.4: 584–603. DOI: 10.1111/flan.12158

Brown, Adam (1988): *Homophones and homographs in Thai, and their implications* (= Forum Phoneticum, 43). Hamburg: Helmut Buske.

Buchmann, Franziska (2015): *Die Wortzeichen im Deutschen* (= Germanistische Bibliothek, 56). Heidelberg: Winter.

Bugarski, Ranko (1993): Graphic relativity and linguistic constructs. In Robert J. Scholes (ed.), *Literacy and language analysis*, 5–18. Hillsdale: Erlbaum.

Bunčić, Daniel, Sandra L. Lippert & Achim Rabus (eds.) (2016): *Biscriptality: a sociolinguistic typology* (= Akademiekonferenzen, 24). Heidelberg: Winter.

Bunčić, Daniel (2016a): A heuristic model for typology. In Daniel Bunčić, Sandra L. Lippert & Achim Rabus (eds.), *Biscriptality: a sociolinguistic typology* (= Akademiekonferenzen, 24), 51–71. Heidelberg: Winter.

Bunčić, Daniel (2016b): Digraphia: medieval Scandinavia, diamesic digraphia. In Daniel Bunčić, Sandra L. Lippert & Achim Rabus (eds.), *Biscriptality: a sociolinguistic typology* (= Akademiekonferenzen, 24), 74–76. Heidelberg: Winter.

Butt, Matthias & Peter Eisenberg (1990): Sprechsilbe und Schreibsilbe. In Christian Stetter (ed.), *Zu einer Theorie der Orthographie: Interdisziplinäre Aspekte gegenwärtiger Schrift- und Orthographieforschung* (= Reihe Germanistische Linguistik, 99), 34–64. Berlin, Boston: De Gruyter. DOI: 10.1515/9783111372280.34

Cahill, Michael (2014): Non-linguistic factors in orthographies. In Michael Cahill & Keren Rice (eds.), *Developing orthographies for unwritten languages*, 9–25. Dallas: SIL International.

Cahill, Michael & Elke Karan (2008): Factors in designing effective orthographies for unwritten languages. *SIL Electronic Working Papers* 2008–001.

Caldwell, Christine A. & Kenny Smith (2012): Cultural evolution and perpetuation of arbitrary communicative conventions in experimental microsocieties. *PLoS ONE* 7.8: e43807. DOI: 10.1371/journal.pone.0043807

Carreiras, Manuel, Manuel Perea, Cristina Gil-López, Reem Abu Mallouh & Elena Salillas (2013): Neural correlates of visual versus abstract letter processing in Roman and Arabic scripts. *Journal of Cognitive Neuroscience* 25.11: 1975–1985. DOI: 10.1162/jocn_a_00438

Castles, Anne & Max Coltheart (2004): Is there a causal link from phonological awareness to success in learning to read? *Cognition* 91.1: 77–111. DOI: 10.1016/S0010-0277(03)00164-1

Catach, Nina (ed.) (1988): *Pour une théorie de la langue écrite. Actes de la Table Ronde internationale C.N.R.S – H.E.S.O. Paris, 23–24 octobre 1986*. Paris: Éditions du Centre National de la Recherche Scientifique.

Cattell, James McK. (1885): Über die Zeit der Erkennung und Benennung von Schriftzeichen, Bildern und Farben. *Philosophische Studien* 2: 635–650.

Chang, Li-Yun, David C. Plaut & Charles A. Perfetti (2016): Visual complexity in orthographic learning: modeling learning across writing system variation. *Scientific Studies of Reading* 20.1: 64–85. DOI: 10.1080/10888438.2015.1104688

Chang, Li-Yun, Yen-Chi Chen & Charles A. Perfetti (2018): GraphCom: a multidimensional measure of graphic complexity applied to 131 written languages. *Behavior Research Methods* 50.1: 427–449. DOI: 10.3758/s13428-017-0881-y

Chang, Ya-Ning, Steve Furber & Stephen Welbourne (2012): Modelling normal and impaired letter recognition: implications for understanding pure alexic reading. Neuropsychologia 50.12: 2773–2788. DOI: 10.1016/j.neuropsychologia.2012.07.031

Changizi, Mark A. & Shinsuke Shimojo (2005): Character complexity and redundancy in writing systems over human history. *Proceedings of the Royal Society B* 272: 267–275. DOI: 10.1098/rspb.2004.2942

Changizi, Mark A., Qiong Zhang, Hao Ye & Shinsuke Shimojo (2006): The structures of letters and symbols throughout human history are selected to match those found in objects in natural scenes. *The American Naturalist* 167.5: E117–E139. DOI: 10.1086/502806

Chao, Yuen R. (1968): *Language and symbolic systems*. Cambridge: CUP.

Chen, Jenn-Yeu & Rong-Ju Cherng (2013): The proximate unit in Chinese handwritten character production. *Frontiers in Psychology* 4: 517. DOI: 10.3389/fpsyg.2013.00517

Chen, Lin (1982): Topological structure in visual perception. *Science* 218.4573: 699–700. DOI: 10.1126/science.7134969

Chen, Ping (2004): *Modern Chinese: history and sociolinguistics*. Cambridge: CUP.

Cherng, Rong-Ju, Chu-Lin Chang & Jenn-Yeu Chen (2009): A New Look at Gender Inequality in Chinese: A Study of Chinese Speakers' Perception of Gender-Based Characters. *Sex Roles* 61: 427–433. DOI: 10.1007/s11199-009-9639-z

Chin, Ng Bee & Kate Burridge (1993): The female radical: portrayal of women in the Chinese script. In Joanne Winter & Gillian Wigglesworth (eds.), *Language and gender in the Australian context* (= Australian Review of Applied Linguistics, Series S, 10), 54–85. Amsterdam, Philadelphia: John Benjamins. DOI: 10.1075/aralss.10.04chi

Chipman, Susan F. (1977): Complexity and structure in visual patterns. *Journal of Experimental Psychology: General* 106.3: 269–301. DOI: 10.1037/0096-3445.106.3.269

Chipman, Susan F. & Morton Mendelson (1979): Influence of six types of visual structure on complexity judgements in children and adults. *Journal of Experimental Psychology: Human Perception & Performance* 5.2: 365–378. DOI: 10.1037/0096-1523.5.2.365

Chou, Siegen K. (1930): Reading and legibility of Chinese characters. II. *Journal of Experimental Psychology* 13.4: 332–351. DOI: 10.1037/h0075030

Christianson, Kiel, Rebecca L. Johnson & Keith Rayner (2005): Letter transpositions within and across morphemes. *Journal of Experimental Psychology* 31.6: 1327–1339. DOI: 10.1037/0278-7393.31.6.1327

Christmann, Ursula (2016): Kognitionspsychologische Ansätze. In Ursula Rautenberg & Ute Schneider (eds.), *Lesen: Ein interdisziplinäres Handbuch*, 21–46. Berlin, Boston: De Gruyter. DOI: 10.1515/9783110275537-004

Clay, Marie M. (1975): *What did I write?* Portsmouth, NH: Heinemann.

Coltheart, Max (1981): Disorders of reading and their implications for models of normal reading. *Visible Language* 15.3: 245–286.

Coltheart, Max, Kathleen Rastle, Conrad Perry, Robyn Langdon & Johannes Ziegler (2001): DRC: a dual route cascaded model of visual word recognition and reading aloud. *Psychological Review* 108.1: 204–256. DOI: 10.1037/0033-295X.108.1.204

Cook, Vivian & Benedetta Bassetti (2005): An introduction to researching second language writing systems. In Vivian Cook & Benedetta Bassetti (eds.), *Second language writing systems* (= Second Language Acquisition, 11), 1–67. Clevedon: Multilingual Matters.

Cooper, Doug (1996): How do Thais tell letters apart? In *Pan-Asiatic linguistics: proceedings of the fourth International Symposium on Language and Linguistics, January 8–10*, 1996, 1163–1176. Salaya: Institute of Language and Culture for Rural Development, Mahidol University at Salaya.

Corder, Stephen P. (1967): The significance of learner's errors. *International Review of Applied Linguistics* 5.4: 161–170. DOI: 10.1515/iral.1967.5.1-4.161

Cossu, Giuseppe, Donald Shankweiler, Isabelle Y. Liberman & Maria Gugliotta (1995): Visual and phonological determinants of misreadings in a transparent orthography. *Reading and Writing* 7.3: 237–256. DOI: 10.1007/BF02539523

Cotter, Katherine N., Paul J. Silvia, Marco Bertamini, Letizia Palumbo & Oshin Vartanian (2017): Curve appeal: exploring individual differences in preference for curved versus angular objects. *i-Perception*. DOI: 10.1177/2041669517693023

Cotticelli-Kurras, Paola & Alfredo Rizza (eds.) (2017): *Variation within and among writing systems: concepts and methods in the analysis of ancient written documents* (= LautSchriftSprache/ScriptandSound, 1). Wiesbaden: Reichert.

Coulmas, Florian (1981): Über Schrift. Frankfurt a. M.: Suhrkamp.

Coulmas, Florian (1989): *The writing systems of the world*. Oxford: Wiley-Blackwell.

Coulmas, Florian (1996a): *The Blackwell encyclopedia of writing systems*. Oxford: Wiley-Blackwell.

Coulmas, Florian (1996b): Typology of writing systems. In Hartmut Günther & Otto Ludwig (eds.), *Schrift und Schriftlichkeit/Writing and its use* (= Handbooks of Linguistics and Communication Science, 10.2), 1380–1387. Berlin, Boston: De Gruyter. DOI: 10.1515/9783110147445.2.9.1380

Coulmas, Florian (2000): The nationalization of writing. *Studies in the Linguistic Sciences* 30.1: 47–59.

Coulmas, Florian (2003): *Writing systems: an introduction to their linguistic analysis* (= Cambridge Textbooks in Linguistics). Cambridge: CUP.

Coulmas, Florian (2009): Evaluating merit – the evolution of writing reconsidered. *Writing Systems Research* 1.1: 5–17. DOI: 10.1093/wsr/wsp001

Coulmas, Florian (2014): Writing systems and language contact in the Euro- and Sinocentric worlds. *Applied Linguistics Review* 5.1: 1–21. DOI: 10.1515/applirev-2014-0001

Coulmas, Florian (2016): Prescriptivism and writing systems. In Ingrid Tieken-Boon van Ostade & Carol Percy (eds.), *Prescription and tradition in language: establishing standards across time and space*, 39–56. Bristol: Multilingual Matters.

Coulmas, Florian & Konrad Ehlich (eds.) (1983): *Writing in focus* (= Trends in Linguistics, 24). Berlin, Boston: De Gruyter. DOI: 10.1515/9783110822830

Cox, Charles H. III, Philippe Coueignoux, Barry A. Blesser & Murray Eden (1982): Skeletons:

A link between theoretical and physical letter descriptions. *Pattern Recognition* 15.1: 11–22. DOI: 10.1016/0031-3203(82)90056-5

Crocco Galèas, Grazia (1998): *The parameters of Natural Morphology*. Padova: Unipress.

Daniels, Peter T. (1990): Fundamentals of grammatology. *Journal of the American Oriental Society* 110.4: 727–731. DOI: 10.2307/602899

Daniels, Peter T. (1991): Is a structural graphemics possible? *LACUS Forum* 18: 528–537.

Daniels, Peter T. (1992): The syllabic origin of writing and the segmental origin of the alphabet. In Pamela Downing, Susan D. Lima & Michael Noonan (eds.), *The linguistics of literacy* (= Typological Studies in Language, 21), 83–110. Amsterdam, Philadelphia: John Benjamins. DOI: 10.1075/tsl.21.10dan

Daniels, Peter T. (1996): The study of writing systems. In Peter T. Daniels & William Bright (eds.), *The world's writing systems*, 3–17. Oxford: OUP.

Daniels, Peter T. (2006a): On beyond alphabets. *Written Language and Literacy* 9.1: 7–24. DOI: 10.1075/wll.9.1.03dan

Daniels, Peter T. (2006b): How many kinds of writing systems are there? In E. M. Rickerson & Barry Hilton (eds.), *The five-minute linguist*, 41–45. London: Equinox.

Daniels, Peter T. (2007): Littera ex occidente: toward a functional history of writing. In Cynthia L. Miller & Charles E. Jones (eds.), *Studies in Semitic and Afro-Asiatic linguistics presented to Gene B. Gragg* (= Studies in Ancient Oriental Civilization, 60), 53–68. Chicago: Oriental Institute of the University of Chicago.

Daniels, Peter T. (2009): Grammatology. In David R. Olson & Nancy Torrance (eds.), *The Cambridge handbook of literacy*, 25–45. Cambridge: CUP.

Daniels, Peter T. (2013): The history of writing as the history of linguistics. In Keith Allan (ed.), *The Oxford handbook of the history of linguistics*, 53–69. Oxford: OUP. DOI: 10.1093/oxfordhb/9780199585847.013.0003

Daniels, Peter T. (2017): Writing systems. In Mark Aronoff & Janie Rees-Miller (eds.), *The handbook of linguistics*, 2nd edition, 75–94. Oxford: Wiley-Blackwell. DOI: 10.1002/9781119072256.ch5

Daniels, Peter T. (2018): *An exploration of writing*. Bristol: Equinox.

Daniels, Peter T. & William Bright (eds.) (1996): *The world's writing systems*. Oxford: OUP.

Daniels, Peter T. & David L. Share (2018): Writing system variation and its consequences for reading and dyslexia. *Scientific Studies of Reading* 22.1: 101–116. DOI: 10.1080/10888438.2017.1379082

Danna, Jérémy, Delphine Massendari, Benjamin Fundari & Stéphanie Ducrot (2018): The optimal viewing position effect in printed versus cursive words: evidence of a reading cost for the cursive font. *Acta Psychologica* 188: 110–121. DOI: 10.1016/j.actpsy.2018.06.003

Danvivathana, Nantana (1987): *The Thai writing system* (= Forum Phoneticum, 39). Hamburg: Helmut Buske.

Davidson, Andrew (2019): Writing: the re-construction of language. *Language Sciences* 72: 134–149. DOI: 10.1016/j.langsci.2018.09.004

Davis, Alyson, Bart De Bruyn & Sam Boyles (2005): Angular figures constrain the perpendicular bias in children's line copying. *Cognitive Development* 20.3: 422–426. DOI: 10.1016/j.cogdev.2005.05.005

Davis, Colin J. (2010): The spatial coding model of visual word identification. *Psychological Review* 117.3: 713–758. DOI: 10.1037/a0019738

de Saussure, Ferdinand (1916): *Cours de linguistique générale*. Paris: Payot.

DeFrancis, John (1943): The alphabetization of Chinese. *Journal of the American Oriental Society* 63.4: 225–240. DOI: 10.2307/594356

DeFrancis, John (1984): *The Chinese language: fact and fantasy*. Honolulu: University of Hawaii Press.

DeFrancis, John (1989): *Visible speech: the diverse oneness of writing systems*. Honolulu: University of Hawaii Press.

Dehaene, Stanislas (2009): *Reading in the brain: the new science of how we read.* London: Penguin.

Derrida, Jacques (1967): *De la grammatologie*. Paris: Ed. de Minuit.

Dervin, Fred (2012): Cultural identity, representation and othering. In Jane Jackson (ed.), *The Routledge handbook of language and intercultural communication*, 181–194. London, New York: Routledge.

Diessel, Holger (2017): Usage-based linguistics. In Mark Aronoff (ed.), *Oxford research encyclopedias: linguistics*. Oxford: OUP. DOI: 10.1093/acrefore/9780199384655.013.363

Ding, Ersu (2005): Hidden iconicity: a Peircean perspective on the Chinese picto-phonetic sign. *Semiotica* 154.1–4: 273–285. DOI: 10.1515/semi.2005.2005.154-1-4.273

Dingemanse, Mark, Damián E. Blasi, Gary Lupyan, Morten H. Christiansen & Padraic Monaghan (2015): Arbitrariness, iconicity, and systematicity in language. *Trends in Cognitive Sciences* 19.10: 603–615. DOI: 10.1016/j.tics.2015.07.013

Diringer, David (1948): *The alphabet: a key to the history of humankind*. London: Hutchinson.

Diringer, David (1962): *Writing*. London: Thames & Hudson.

Dokumente zur Interpunktion europäischer Sprachen (1939): Gothenburg.

Domahs, Ulrike & Beatrice Primus (2015): Laut – Gebärde – Buchstabe. In Ekkehard Felder & Andreas Gardt (eds.), *Handbuch Sprache und Wissen* (= Handbooks of Linguistic Knowledge, 1), 125–142. Berlin, Boston: De Gruyter. DOI: 10.1515/9783110295979.125

Domahs, Ulrike & Beatrice Primus (eds.) (2016): *Handbuch Laut, Gebärde, Buchstabe* (= Handbooks of Linguistic Knowledge, 2). Berlin, Boston: De Gruyter. DOI: 10.1515/9783110295993

Donegan, Patricia J. [1978] (1985): *On the Natural Phonology of vowels.* PhD Dissertation, Ohio State University. (Published 1985 in Ohio State University Working Papers in Linguistics 23.)

Donegan, Patricia & David Stampe (1979): The study of Natural Phonology. In Daniel A. Dinnsen & Stephen R. Anderson (eds.), *Current approaches to phonological theory*, 126–173. Bloomington: Indiana University Press.

Donegan, Patricia & David Stampe (2009): Hypotheses of Natural Phonology. *Poznań Studies in Contemporary Linguistics* 45.1: 1–31. DOI: 10.2478/v10010-009-0002-x

Downey, Greg (2014): All forms of writing. *Mind & Language* 29.3: 304–319. DOI: 10.1111/mila.12052

Dresher, B. Elan (2011): The phoneme. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell companion to phonology,* vol. 1: *General issues and segmental phonology*, 241–266. Oxford: Wiley-Blackwell.

Dressler, Wolfgang U. (1980): Naturalness as a principle in genetic and typological linguistics: introduction. In Torben Thrane, Vibeke Winge, Lachlan Mackenzie, Una Canger & Niels Ege (eds.), *Typology and genetics of language* (= Travaux du cercle linguistique de Copenhague, XX), 75–91. Copenhagen: Linguistic Circle of Copenhagen.

Dressler, Wolfgang U. (1984): Explaining Natural Phonology. *Phonology* 1: 29–51. DOI: 10.1017/S0952675700000282

Dressler, Wolfgang U. (1987): Semiotische Grundlagen einer Theorie der Natürlichen Phonologie und Morphologie. In Jeff Bernard (ed.), *Semiotica Austriaca. Angewandte Semiotik 9, 10*, 165–172. Wien: Österreichische Gesellschaft für Semiotik.

Dressler, Wolfgang U. (1989): *Semiotische Parameter einer textlinguistischen Natürlichkeitstheo-*

*rie* (= Veröffentlichungen der Kommission für Linguistik und Kommunikationsforschung, 21). Wien: Verlag der Österreichischen Akademie der Wissenschaften.

Dressler, Wolfgang U. (1990): The cognitive perspective of 'naturalist' linguistic models. *Cognitive Linguistics* 1.1: 75–98. DOI: 10.1515/cogl.1990.1.1.75

Dressler, Wolfgang U. (1999): On a semiotic theory of preferences in language. *Peirce Seminar Papers* 4: 389–415.

Dressler, Wolfgang U. (2000): Naturalness. In Geert Booij, Christian Lehmann, Joachim Mugdan, Wolfgang Kesselheim & Stavros Skopeteas (eds.), *Morphology: an international handbook of inflection and word-formation*, vol. 1. (= Handbooks of Linguistics and Communication Science, 17.1), 288–296. Berlin, Boston: De Gruyter. DOI: 10.1515/9783110111286.1.4.288

Dressler, Wolfgang U. (2006): Natural Morphology. In Keith Brown (ed.), *Encyclopedia of language and linguistics*, 539–540. Amsterdam: Elsevier.

Dressler, Wolfgang U., Willi Mayerthaler, Oswald Panagl & Wolfgang U. Wurzel (1987): *Leitmotifs in Natural Morphology* (= Studies in Language Companion Series, 10). Amsterdam, Philadelphia: John Benjamins. DOI: 10.1075/slcs.10

Dressler, Wolfgang U. & Marianne Kilani-Schoch (2016): Natural Morphology. In Andrew Hippisley & Gregory Stump (eds.), *The Cambridge handbook of morphology*, 356–389. Cambridge: CUP.

Drews, Catharina (2011): *Werden längenhaltige Wörter schneller gelesen als längenlose?* Master's thesis, Universität Oldenburg. <https://www.uni-oldenburg.de/fileadmin/user_upload/germanistik/personen/nanna.fuhrhop/Drews_Lesen_laengenhaltiger_und_laengenloser_Woerter.pdf> (September 20th, 2020).

Dryer, Matthew S. (2006): Descriptive theories, explanatory theories, and Basic Linguistic Theory. In Felix K. Ameka, Alan Dench & Nicholas Evans (eds.), *Catching language – The standing challenge of grammar writing*, 207–234. Berlin, Boston: De Gruyter. DOI: 10.1515/9783110197693.207

Duden (2017): *Duden. Die deutsche Rechtschreibung*, 27th edition. Berlin: Duden.

Duden (2020): *Duden. Die deutsche Rechtschreibung*, 28th edition. Berlin: Duden.

Duñabeitia, Jon A., Manuel Perea & Manuel Carreiras (2007): Do transposed-letter similarity effects occur at a morpheme level? Evidence for morpho-orthographic decomposition. *Cognition* 105.3: 691–703. DOI: 10.1016/j.cognition.2006.12.001

Dürscheid, Christa (2000): Verschriftungstendenzen jenseits der Rechtschreibreform. *Zeitschrift für germanistische Linguistik* 28: 237–247. DOI: 10.1515/zfgl.2000.28.2.223

Dürscheid, Christa (2016): *Einführung in die Schriftlinguistik*, 5th edition. Göttingen: Vandenhoeck & Ruprecht.

Dürscheid, Christa (2018): Bild, Schrift, Unicode. In Jürgen Hermes, Jean-Yves Lalande, Guido Mensching & Claes Neuefeind (eds.), *Sprache – Mensch – Maschine. Beiträge zu Sprache und Sprachwissenschaft, Computerlinguistik und Informationstechnologie für Jürgen Rolshoven aus Anlass seines sechsundsechzigsten Geburtstages*, 269–285. Köln: Universität zu Köln.

Dziubalska-Kołaczyk, Katarzyna (2002): Challenges for Natural Linguistics in the 21st century: a personal view. In Katarzyna Dziubalska-Kołaczyk & Jarosław Weckwerth (eds.), *Future challenges for Natural Linguistics* (= LINCOM Studies in Theoretical Linguistics, 30), 103–128. München: LINCOM.

Edwards, Mary L. & Lawrence D. Shriberg (1983): *Phonology: applications in communicative disorders*. San Diego: College-Hill Press.

Ehlich, Konrad (2001): Graphemics/[Transindividual] graphology. In Rudolf de Cillia, Hans-Jürgen Krumm & Ruth Wodak (eds.), *Loss of communication in the information age*, 61–76. Wien: Verlag der Österreichischen Akademie der Wissenschaften.

Ehlich, Konrad, Florian Coulmas & Gabriele Graefen (eds.) (1996): *A bibliography on writing and written language* (= Trends in Linguistics, 89). Berlin, Boston: De Gruyter. DOI: 10.1515/9783110889352

Eira, Christina (1998): Authority and discourse: towards a model for orthography selection. *Written Language and Literacy* 1.2: 171–224. DOI: 10.1075/wll.1.2.03eir

Eisenberg, Peter (1988): Über die Autonomie der graphematischen Analyse. In Dieter Nerius & Gerhard Augst (eds.), *Probleme der geschriebenen Sprache. Beiträge zur Schriftlinguistik auf dem XIV. internationalen Linguistenkongreß 1987 in Berlin* (= Linguistische Studien, A, 173), 25–35. Berlin: Akademie der Wissenschaften der DDR.

Eisenberg, Peter (1989): Die Schreibsilbe im Deutschen. In Peter Eisenberg & Hartmut Günther (eds.), *Schriftsystem und Orthographie* (= Reihe Germanistische Linguistik, 97), 57–84. Berlin, Boston: De Gruyter. DOI: 10.1515/9783111372266.57

Eisenberg, Peter (2006): *Grundriss der deutschen Grammatik. Bd. 1: Das Wort.* 3rd, revised edition. Stuttgart: Metzler.

Eisenberg, Peter (2013): *Grundriss der deutschen Grammatik. Bd. 1: Das Wort*. 4th, revised edition. Stuttgart: Metzler.

Eisenberg, Peter (2017): *Deutsche Orthografie: Regelwerk und Kommentar*. Berlin, Boston: De Gruyter. DOI: 10.1515/9783110525229

Ellis, Nick C., Miwa Natsume, Katerina Stavropoulou, Lorenc Hoxhallari, Victor H. P. van Daal, Nicoletta Polyzoe, Maria-Louisa Tsipa & Michalis Petalas (2004): The effects of orthographic depth on learning to read alphabetic, syllabic, and logographic scripts. *Reading Research Quarterly* 39.4: 438–468. DOI: 10.1598/RRQ.39.4.5

Elmentaler, Michael (2018): *Historische Graphematik des Deutschen: Eine Einführung* (= narr Studienbücher). Tübingen: Narr.

Elsen, Hilke (2014): *Linguistische Theorien* (= narr Studienbücher). Tübingen: Narr.

Elson, Benjamin & Velma Pickett (1962): *An introduction to morphology and syntax*. Santa Ana: Summer Institute of Linguistics.

Enderle, Ursula (2005): *Autonomie der geschriebenen Sprache? Zur Theorie phonographischer Beschreibungskategorien am Beispiel des Deutschen* (= Philologische Studien und Quellen, 188). Berlin: Erich Schmidt Verlag.

Evertz, Martin (2016): Graphematischer Fuß und graphematisches Wort. In Ulrike Domahs & Beatrice Primus (eds.), *Handbuch Laut, Gebärde, Buchstabe* (= Handbooks of Linguistic Knowledge, 2), 377–397. Berlin, Boston: De Gruyter. DOI: 10.1515/9783110295993-021

Evertz, Martin (2018): *Visual prosody: the graphematic foot in English and German*. Berlin, Boston: De Gruyter. DOI: 10.1515/9783110583441

Eviatar, Zohar & Raphiq Ibrahim (2004): Morphological and orthographic effects on hemispheric. *Reading and Writing* 17.7–8: 691–705. DOI: 10.1007/s11145-004-2659-8

Ewald, Petra (2007): Zur orthographischen Regel. In Dieter Nerius (et al.) (eds.), *Deutsche Orthographie*, 4th edition, 40–54. Hildesheim: Olms.

Faber, Alice (1992): Phonemic segmentation as epiphenomenon: evidence from the history of alphabetic writing. In Pamela Downing, Susan D. Lima & Michael Noonan (eds.), *The linguistics of literacy* (= Typological Studies in Language, 21), 111–134. Amsterdam: John Benjamins. DOI: 10.1075/tsl.21.11fab

Fasold, Ralph (1984): *The sociolinguistics of society*. Oxford: Wiley-Blackwell.

Fay, Nicolas, Simon Garrod, Leo Roberts & Nik Swoboda (2010): The interactive evolution of human communication systems. *Cognitive Science* 34.3: 351–386. DOI: 10.1111/j.1551-6709.2009.01090.x

Feldman, Laurie B. & Witina W. T. Siok (1999): Semantic radicals in phonetic compounds: implications for visual character recognition in Chinese. In Jian Wang, Albrecht W.

Inhoff & Hsuan-Chih Chen (eds), *Reading Chinese script: a cognitive analysis*, 19–35. Mahwah: Lawrence Erlbaum.

Fenk-Oczlon, Gertraud & August Fenk (1995): Selbstorganisation und natürliche Typologie. *Sprachtypologie und Universalienforschung* 48.3: 223–238. DOI: 10.1524/stuf.1995.48.3.223

Finkbeiner, Matthew & Max Coltheart (2009): Letter recognition: from perception to representation. *Cognitive Neuropsychology* 26.1: 1–6. DOI: 10.1080/02643290902905294

Fischer, Jean-Paul & Anne-Marie Koch (2016): Mirror writing in 5- to 6-year-old children: the preferred hand is not the explanation. *Laterality: Asymmetries of Body, Brain and Cognition* 21.1: 34–49. DOI: 10.1080/1357650X.2015.1066383

Fischer, Jean-Paul (2017): Character reversal in children: the prominent role of writing direction. *Reading and Writing* 30.3: 523–542. DOI: 10.1007/s11145-016-9688-y

Fiset, Daniel, Caroline Blais, Catherine Éthier-Majcher, Martin Arguin, Daniel Bub & Frédéric Gosselin (2008): Features for identification of uppercase and lowercase letters. *Psychological Science* 19.11: 1161–1168. DOI: 10.1111/j.1467-9280.2008.02218.x

Fiset, Daniel, Caroline Blais, Martin Arguin, Karine Tadros, Catherine Éthier-Majcher, Daniel Bub & Frédéric Gosselin (2009): The spatio-temporal dynamics of visual letter recognition. *Cognitive Neuropsychology* 26.1: 23–35. DOI: 10.1080/02643290802421160

Fisher, Dennis F. (1976): Spatial factors in reading and search: the case for space. In Richard A. Monty & John W. Senders (eds.), *Eye movements and psychological processes*, 417–427. Hillsdale: Lawrence Erlbaum.

Fishman, Joshua (ed.) (1977): *Advances in the creation and revision of writing systems* (= Contributions to the Sociology of Language, 8). Berlin, Boston: De Gruyter. DOI: 10.1515/9783110807097

Flores D'Arcaïs, Giovanni B. (1994): Order of strokes writing as a cue for retrieval in reading Chinese characters. *European Journal of Cognitive Psychology* 6.4: 337–355. DOI: 10.1080/09541449408406519

Fludernik, Monika (1996): Comment on Wolfgang U. Dressler's "Natural Textlinguistics". *Sprachtypologie und Universalienforschung* 49.3: 322–329. DOI: 10.1524/stuf.1996.49.3.322

Fox, Anthony (2000): *Prosodic features and prosodic structure: the phonology of suprasegmentals*. Oxford: OUP.

Frangou, Satu-Maarit, Heli Ruokamo, Riina Parviainen & Jan Wikgren (2018): Can you put your finger on it? The effects of writing modality on Finnish students' recollection. *Writing Systems Research* 10.2: 82–94. DOI: 10.1080/17586801.2018.1536015

Friedman, Rhonda B. (1980): Identity without form: abstract representations of letters. *Perception & Psychophysics* 28.1: 53–60. DOI: 10.3758/BF03204315

Friedmann, Naama & Manar Haddad-Hanna (2012): Letter position dyslexia in Arabic: from form to position. *Behavioural Neurology* 25.3: 193–203. DOI: 10.1155/2012/296974

Frost, Ram (2012): Towards a universal model of reading. *Behavioral and Brain Sciences* 35.5: 263–279. DOI: 10.1017/S0140525X11001841

Fuhrhop, Nanna (n. d.): Die Ikonizität in der Schrift: Zur Form von Minuskeln und Majuskeln.

Fuhrhop, Nanna (2008): Das graphematische Wort (im Deutschen): Eine erste Annäherung. *Zeitschrift für Sprachwissenschaft* 27.2: 189–228. DOI: 10.1515/zfsw.2008.010

Fuhrhop, Nanna (2015): *Orthografie*, 4th edition (= Kurze Einführungen in die germanistische Linguistik, 1). Heidelberg: Winter.

Fuhrhop, Nanna (2018): Graphematik des Deutschen im europäischen Vergleich. In Angelika Wöllstein, Peter Gallmann, Mechthild Habermann & Manfred Krifka (eds.), *Grammatiktheorie und Empirie in der germanistischen Linguistik* (= Germanistische

Sprachwissenschaft um 2020, 1), 587–616. Berlin, Boston: De Gruyter. DOI: 10.1515/9783110490992-020

Fuhrhop, Nanna & Franziska Buchmann (2009): Die Längenhierarchie: Zum Bau der graphematischen Silbe. *Linguistische Berichte* 218: 127–155.

Fuhrhop, Nanna & Franziska Buchmann (2016): Graphematische Silbe. In Ulrike Domahs & Beatrice Primus (eds.), *Handbuch Laut, Gebärde, Buchstabe* (= Handbooks of Linguistic Knowledge, 2), 356–376. Berlin, Boston: De Gruyter. DOI: 10.1515/9783110295993-020

Fuhrhop, Nanna, Franziska Buchmann & Kristian Berg (2011): The length hierarchy and the graphematic syllable. *Written Language and Literacy* 14.2: 275–292. DOI: 10.1075/wll.14.2.05fuh

Fuhrhop, Nanna & Jörg Peters (2013): *Einführung in die Phonologie und Graphematik*. Stuttgart: Metzler.

Fuhrhop, Nanna & Karsten Schmidt (2014): Die zunehmende Profilierung der Schreibsilbe in der Geschichte des Deutschen. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 136.4: 538–568. DOI: 10.1515/bgsl-2014-0047

Fuhrhop, Nanna & Niklas Schreiber (2016): Graphematische Lesehilfen. *Mitteilungen des Deutschen Germanistenverbandes* 63.2: 129–146. DOI: 10.14220/mdge.2016.63.2.129

Fuhrhop, Nanna, Rebecca Carroll, Catharina Drews & Esther Ruigendijk (2016): Sind die Buchstabenformen eine Lesehilfe? *Mitteilungen des Deutschen Germanistenverbandes* 63.2: 119–128. DOI: 10.14220/mdge.2016.63.2.119

Gabler, Birgit (1983): *Untersuchungen zur Variabilität der Orthographie der deutschen Gegenwartssprache*. PhD dissertation, Universität Rostock.

Gaeta, Livio (2006): How to live naturally and not be bothered by economy. *Folia Linguistica* 40.1–2: 7–28. DOI: 10.1515/flin.40.1-2.7

Galambos, Imre (2015): Variant characters. In Rint Sybesma (ed.), *Encyclopedia of Chinese language and linguistics*. Leiden: Brill. DOI: 10.1163/2210-7363_ecll_COM_00000438

Gallmann, Peter (1985): *Graphische Elemente der geschriebenen Sprache: Grundlagen für eine Reform der Orthographie* (= Reihe Germanistische Linguistik, 60). Berlin, Boston: De Gruyter. DOI: 10.1515/9783111630380

Gallmann, Peter (1986): The graphic elements of German written language. In Gerhard Augst (ed.), *New trends in graphemics and orthography*, 43–79. Berlin, Boston: De Gruyter. DOI: 10.1515/9783110867329.43

Gallmann, Peter & Horst Sitta (1997): Zum Begriff der orthographischen Regel. In Gerhard Augst, Karl Blüml, Dieter Nerius & Horst Sitta (eds.), *Die Neuregelung der deutschen Orthographie. Begründung und Kritik* (= Reihe Germanistische Linguistik, 179). 93–110. Berlin, Boston: De Gruyter. DOI: 10.1515/9783110927993.93

Ganopole, Selina J. (1987): The development of word consciousness prior to first grade. *Journal of Reading Behavior* 19.4: 415–436. DOI: 10.1080/10862968709547614

Garbe, Burckhard (1985): Graphemtheorien und mögliche strukturmodelle zur beschreibung der orthographie. In Gerhard Augst (ed.), *Graphematik und Orthographie. Neuere Forschungen der Linguistik, Psychologie und Didaktik in der Bundesrepublik Deutschland*, 1–21. Frankfurt a. M.: Peter Lang.

Garbe, Burckhard (2000): Phonetik und Phonologie, Graphetik und Graphemik des Neuhochdeutschen seit dem 17. Jahrhundert. In Werner Besch, Anne Betten, Oskar Reichmann & Stefan Sonderegger (eds.), *Sprachgeschichte*, 2. Teilband (= Handbooks of Linguistics and Communication Science, 2.1), 1765–1782. Berlin, Boston: De Gruyter. DOI: 10.1515/9783110158823.2.13.1765

Garrod, Simon, Nicholas Fay, John Lee, Jon Oberlander & Tracy McLeod (2007): Foundations of representation: where might graphical symbol systems come from? *Cognitive Science* 31.6: 961–987. DOI: 10.1080/03640210701703659

Gartus, Andreas & Helmut Leder (2017): Predicting perceived visual complexity of abstract patterns using computational measures: the influence of mirror symmetry on complexity perception. *PLoS One* 12.11: e0185276. DOI: 10.1371/journal.pone.0185276

Gauthier, Isabel, Alan C.-N. Wong, William G. Hayward & Olivia S. Cheung (2006): Font tuning associated with expertise in letter perception. *Perception* 35.4: 541–559. DOI: 10.1068/p5313

Gee, James P. (2008): *Social linguistics and literacies: ideology in discourses*, 3rd edition. London: Routledge.

Gelb, Ignace J. [1952] (1969): *A Study of Writing*, 2nd edition (3rd impression). Chicago: University of Chicago Press.

Genesee, Fred, Esther Geva, Cheryl Dressler & Michael L. Kamil (2006): Synthesis: cross-linguistic relationships. In Diane August & Timothy Shanahan (eds.), *Developing literacy in second-language learners: report of the National Literacy Panel on language-minority children and youth*, 153–174. Mahwah, NJ: Lawrence Erlbaum Associates.

Gentner, Donald R. (1983): The acquisition of typewriting skill. *Acta Psychologica* 54.1–3: 233–248. DOI: 10.1016/0001-6918(83)90037-9

Gerth, Sabrina, Thomas Dolk, Annegret Klassert, Michael Fliesser, Martin H. Fischer, Guido Nottbusch & Julia Festman (2016a): Adapting to the surface: a comparison of handwriting measures when writing on a tablet computer and on paper. *Human Movement Science* 48: 62–73. DOI: 10.1016/j.humov.2016.04.006

Gerth, Sabrina, Annegret Klassert, Thomas Dolk, Michael Fliesser, Martin H. Fischer, Guido Nottbusch & Julia Festman (2016b): Is handwriting performance affected by the writing surface? Comparing preschoolers', second graders', and adults' writing performance on a tablet vs. paper. *Frontiers in Psychology* 7: 1308. DOI: 10.3389/fpsyg.2016.01308

Geva, Esther, Lesly Wade-Woolley & Michal Shany (1993): The concurrent development of spelling and decoding in two different orthographies. *Journal of Reading Behavior* 25.4: 383–406. DOI: 10.1080/10862969309547827

Giannouli, Vaitsa (2013): Visual symmetry perception. *Encephalos* 50.1: 31–42.

Gibson, Eleanor J. & Harry Levin (1980): *Die Psychologie des Lesens*. Stuttgart: Klett-Cotta.

Gibson, Eleanor J., Harry Osser, William Schiff & Jesse Smith (1963): An analysis of critical features of letters, tested by a confusion matrix. In *Cooperative Research Project No. 639: a basic research program on reading*. Washington: U.S. Office of Education.

Giese, Heinz W. (1993): ‚Geschriebene Sprache' 1981–1991. Eine Zusammenstellung von Arbeiten der Studiengruppe ‚Geschriebene Sprache' bei der Werner Reimers Stiftung, Bad Homburg. In Jürgen Baurmann, Hartmut Günther & Ulrich Knoop (eds.), *homo scribens. Perspektiven der Schriftlichkeitsforschung* (= Reihe Germanistische Linguistik, 134), 379–396. Berlin, Boston: De Gruyter. DOI: 10.1515/9783111377087.379

Gingras, Maxime & Monique Sénéchal (2019): Evidence of statistical learning of orthographic representations in grades 1–5: the case of silent letters and double consonants in French. *Scientific Studies in Reading* 23.1: 37–48. DOI: 10.1080/10888438.2018.1482303

Glück, Helmut (1987) *Schrift und Schriftlichkeit. Eine sprach- und kulturwissenschaftliche Studie*. Stuttgart: Metzler.

Glück, Helmut (2011): Ikonizität der Schrift? In Antonio Loprieno, Carsten K. Salis & Birgit Mersmann (eds.), *Bild, Macht, Schrift: Schriftkulturen in bildkritischer Perspektive*, 67–79. Weilerswist: Velbrück Wissenschaft.

Glück, Helmut (2016a): Interdependenzhypothese. In Helmut Glück & Michael Rödel (eds.), *Metzler Lexikon Sprache*, 5th edition, 301–302. Stuttgart: Metzler.

Glück, Helmut (2016b): Graphem. In Helmut Glück & Michael Rödel (eds.), *Metzler Lexikon Sprache*, 5th edition, 252–253. Stuttgart: Metzler.

Glück, Helmut (2016c): Schrift. In Helmut Glück & Michael Rödel (eds.), *Metzler Lexikon Sprache*, 5th edition, 593–594. Stuttgart: Metzler.

Glück, Helmut (2016d): Schriftlinguistik. In Helmut Glück & Michael Rödel (eds.), *Metzler Lexikon Sprache*, 5th edition, 596. Stuttgart: Metzler.

Glück, Helmut (2016e): Graphematik. In Helmut Glück & Michael Rödel (eds.), *Metzler Lexikon Sprache*, 5th edition, 253. Stuttgart: Metzler.

Gnanadesikan, Amalia E. (2017): Towards a typology of phonemic scripts. *Writing Systems Research* 9.1: 14–35. DOI: 10.1080/17586801.2017.1308239

Goldin-Meadow, Susan & Rachel I. Mayberry (2001): How do profoundly deaf children learn to read? *Learning Disabilities Research & Practice* 16.4: 222–229. DOI: 10.1111/0938-8982.00022

Gombert, Jean E. & Michel Fayol (1992): Writing in preliterate children. *Learning and Instruction* 2.1: 23–41. DOI: 10.1016/0959-4752(92)90003-5

Gomez, Pablo, Roger Ratcliff & Manuel Perea (2008): The overlap model: a model of letter position coding. *Psychological Review* 115.3: 577–600. DOI: 10.1037/a0012667

Gong, Yushu (2006): Graph typology of ancient Chinese and Sumerian writing systems – a comparative perspective. *Oriental Studies*, Special Issue 6: 39–108.

Goodnow, Jacqueline J. & Rochelle A. Levine (1973): "The grammar of action": sequence and syntax in children's copying. *Cognitive Psychology* 4.1: 82–98. DOI: 10.1016/0010-0285(73)90005-4

Goody, Jack & Ian P. Watt (1963): The consequences of literacy. *Comparative Studies in History and Society* 5.3: 304–345. DOI: 10.1017/S0010417500001730

Gordon, Jurij (2006): *Kniga pro bukvy ot Aa do Jaja*. Moscow: Izdatel'stvo Studii Artemija Lebedeva.

Gosse, Claire, Simon Carbonnelle, Christophe de Vleeschouwer & Marie Van Reybroeck (2018): Specifying the graphic characteristics of words that influence children's handwriting. *Reading and Writing* 31.5: 1181–1207. DOI: 10.1007/s11145-018-9834-9

Grainger, Jonathan & Walter J. B. Van Heuven (2004): Modeling letter position coding in printed word perception. In Patrick Bonin (ed.), *Mental lexicon: "some words to talk about words"*, 1–24. New York: Nova Science.

Gredig, Andi (2020): *Schreiben mit der Hand. Begriffe – Diskurs – Praktiken*. PhD dissertation, Universität Zürich.

Gregg, Noël (1995): *Written expression disorders* (= Neuropsychology and Cognition, 10). Dordrecht: Springer. DOI: 10.1007/978-94-011-0297-1

Grenoble, Lenore A. & Lindsay J. Whaley (2006): *Saving languages: an introduction to language revitalization*. Cambridge: CUP.

Gronemeyer, Sven (2015): Class struggle: Towards a better understanding of Maya writing using comparative graphematics. In Harri Kettunen & Christophe Helmke (eds.), *On methods: How we know what we think we know about the Maya. Proceedings of the 17th European Maya Conference, Helsinki December 9–15, 2012* (= Acta Mesoamericana, 28), 101–117. München: Anton Saurwein.

Grosswiler, Paul (2004): Dispelling the Alphabet Effect. *Canadian Journal of Communication* 29.2: 145–158.

Gruendler, Beatrice (2012): Stability and change in Arabic script. In Stephen D. Houston (ed.), *The shape of script: how and why writing systems change*, 93–118. Santa Fe: School for Advanced Research Press.

Günther, Hartmut (1988): *Schriftliche Sprache: Strukturen geschriebener Wörter und ihre Verar-*

*beitung beim Lesen* (= Konzepte der Sprach- und Literaturwissenschaft, 40). Berlin, Boston: De Gruyter. DOI: 10.1515/9783110935851

Günther, Hartmut (1990a): Zur neueren Schriftlichkeitsforschung. *Beiträge zur Geschichte der deutschen Sprache* 112: 349–370. DOI: 10.1515/bgsl.1990.1990.112.349

Günther, Hartmut (1990b): Typographie, Orthographie, Graphetik. Überlegungen zu einem Buch von Otl Aicher. In Christian Stetter (ed.), *Zu einer Theorie der Orthographie. Interdisziplinäre Aspekte gegenwärtiger Schrift- und Orthographieforschung* (= Reihe Germanistische Linguistik, 99), 90–103. Berlin, Boston: De Gruyter. DOI: 10.1515/9783111372280.90

Günther, Hartmut (1993a): Die Studiengruppe ‚Geschriebene Sprache' bei der Werner Reimers Stiftung, Bad Homburg. In Jürgen Baurmann, Hartmut Günther & Ulrich Knoop (eds.), *homo scribens. Perspektiven der Schriftlichkeitsforschung* (= Reihe Germanistische Linguistik, 134), 371–378. Berlin, Boston: De Gruyter. DOI: 10.1515/9783111377087.371

Günther, Hartmut (1993b): Graphetik – Ein Entwurf. In Jürgen Baurmann, Hartmut Günther & Ulrich Knoop (eds.), *homo scribens. Perspektiven der Schriftlichkeitsforschung* (= Reihe Germanistische Linguistik, 134), 29–42. Berlin, Boston: De Gruyter. DOI: 10.1515/9783111377087.29

Günther, Hartmut & Otto Ludwig (eds.) (1994/1996): *Schrift und Schriftlichkeit/Writing and its use* (= Handbooks of Linguistics and Communication Science, 10.1 & 10.2). Berlin, Boston: De Gruyter.

Güthert, Kerstin (2016): *Bericht des Rats für deutsche Rechtschreibung über die Wahrnehmung seiner Aufgaben in der Periode 2011 bis 2016.* Mannheim: Rat für deutsche Rechtschreibung. <http://www.rechtschreibrat.com/DOX/rfdr_Bericht_2011–2016.pdf> (January 19th, 2019).

Haas, Mary R. (1956): *The Thai system of writing*. Washington: American Council of Learned Societies.

Häffner, Sonja (2009): *Universalien der Schrift und Schriftverarbeitung: zu den schrifttheoretischen Prämissen psychologischer Lesemodelle* (= Europäische Hochschulschriften 21, Linguistik 349). Frankfurt a. M.: Peter Lang.

Häffner, Sonja (2013): Eszett. In Martin Neef, Said Sahel & Rüdiger Weingarten (eds.), *Schriftlinguistik/Grapholinguistics* (= Dictionaries of Linguistics and Communication Science, 5). Berlin, Boston: De Gruyter. DOI: 10.1515/wsk.5.0.eszett

Halle, Morris (1959): *The sound pattern of Russian: a linguistic an acoustical investigation* (= Description and Analysis of Contemporary Standard Russian, 1). Berlin, Boston: De Gruyter. DOI: 10.1515/9783110869453

Halliday, Michael A. K. [1977] (2010): Ideas about language. In Jonathan J. Webster & M. A. K. Halliday (eds.), *On language and linguistics*, 92–112. London, New York: Continuum.

Hamp, Eric P. (1959): Graphemics and paragraphemics. *Studies in Linguistics* 14.1–2: 1–5.

Handel, Zev (2013): Can a logographic script be simplified? Lessons from the 20th century Chinese writing reform informed by recent psycholinguistic research. *Scripta* 5: 21–66.

Hannas, William C. (1997): *Asia's orthographic dilemma*. Honolulu: University of Hawai'i Press.

Haralambous, Yannis (2007): *Fonts & Encodings*. Sebastopol, CA: O'Reilly.

Haralambous, Yannis (2020): Grapholinguistics, TEX, and a June 2020 conference in Paris. *TUGboat* 41.1: 12–19.

Haralambous, Yannis & Joseph Dichy (2019): Graphemic methods for gender-neutral writing. In Yannis Haralambous (ed.), *Graphemics in the 21st century. Brest, June 13–15,* *2018. Procedings* (= Grapholinguistics and Its Applications, 1), 41–89. Brest: Fluxus Editions. DOI: 10.36824/2018–graf–hara2

Harralson, Heidi H. (2013): *Developments in handwriting and signature identification in the digital age*. Oxford, Waltham: Anderson Publishing.

Harris, Roy (1980): *The language-makers*. London: Duckworth.

Harris, Roy (1994): Semiotic aspects of writing. In Hartmut Günther & Otto Ludwig (eds.), *Schrift und Schriftlichkeit/Writing and its use* (= Handbooks of Linguistics and Communication Science, 10.1), 41–48. Berlin, Boston: De Gruyter. DOI: 10.1515/9783110111293.1.1.41

Harris, Roy (2005): Schrift und linguistische Theorie. In Gernot Grube, Werner Kogge & Sybille Krämer (eds.), *Schrift: Kulturtechnik zwischen Auge, Hand und Maschine* (= Kulturtechnik), 61–80. München: Wilhelm Fink.

Haspelmath, Martin (2010): Comparative concepts and descriptive categories in crosslinguistic studies. *Language* 86.3: 663–687. DOI: 10.1353/lan.2010.0021

Hatcher, Lynley (2008): Script change in Azerbaijan: acts of identity. *International Journal of the Sociology of Language* 192: 105–116. DOI: 10.1515/IJSL.2008.038

Heller, Klaus (1980): Zum Graphembegriff. In Dieter Nerius & Jürgen Scharnhorst (eds.), *Theoretische Probleme der deutschen Orthographie* (= Sprache und Gesellschaft, 16), 74–108. Berlin: Akademie-Verlag.

Hennig, Mathilde (2006): *Grammatik der gesprochenen Sprache in Theorie und Praxis*. Kassel: kassel university press.

Hermena, Ehab W., Mercedes Sheen, Maryam Al Jassmi, Khulood Al Falasi, Maha Al Matroushi & Timothy R. Jordan (2017): Reading rate and comprehension for text presented on tablet and paper: evidence from Arabic. *Frontiers in Psychology* 8: 257. DOI: 10.3389/fpsyg.2017.00257

Herrick, Earl M. (1966): *A linguistic description of Roman alphabets* (= Hartford Studies in Linguistics, 19). Hartford: Hartford Seminary Foundation.

Herrick, Earl M. (1974): A taxonomy of alphabets and scripts. *Visible Language* 8.1: 5–32.

Herrick, Earl M. (1975a): Orthographies and scripts: a necessary distinction. *LACUS Forum* 1: 533–539.

Herrick, Earl M. (1975b): Letters with alternative basic shapes. *Visible Language* 9.2: 133–144.

Herrick, Earl M. (1994): Of course a structural graphemics is possible! *LACUS Forum* 21: 413–424.

Hill, Archibald A. (1967): The typology of writing systems. In William M. Austin (ed.), *Papers in linguistics in honor of Léon Dostert*, 92–99. Berlin, Boston: De Gruyter. DOI: 10.1515/9783111675886-008

Hill, Robin L. & Wayne S. Murray (2000): Commas and spaces: effects of punctuation on eye movements and sentence parsing. In Alan Kennedy, Ralph Radach, Dieter Heller & Joël Pynte (eds.), *Reading as a perceptual process*, 565–596. Amsterdam: Elsevier. DOI: 10.1016/B978-008043642-5/50027-9

Hinton, Leanne (2014): Orthography wars. In Michael Cahill & Keren Rice (eds.), *Developing orthographies for unwritten languages*, 139–168. Dallas: SIL International.

Hippisley, Andrew, Marina Chumakina, Greville G. Corbett & Dunstan Brown (2004): Suppletion: frequency, categories and distribution of stems. *Studies in Language* 28.2: 387–418. DOI: 10.1075/sl.28.2.05hip

Ho, Connie S.-H. & Peter Bryant (1997): Learning to read Chinese beyond the logographic phase. *Reading Research Quarterly* 32.3: 276–289. DOI: 10.1598/RRQ.32.3.3

Ho, Connie S.-H., Ting-Ting Ng & Wing-Kin Ng (2003): A 'radical' approach to reading development in Chinese: the role of semantic radicals and phonetic radicals. *Journal of Literacy Research* 35.3: 849–878. DOI: 10.1207/s15548430jlr3503_3

Hockett, Charles F. (1951): Review of *Nationalism and Language Reform in China*, by J. DeFrancis. *Language* 27.3: 439–445. DOI: 10.2307/409788

Hockett, Charles F. (1960): The origin of speech. *Scientific American* 203.3: 88–96. DOI: 10.1038/scientificamerican0960-88

Hoffmann-Dilloway, Erika (2013): (Don't) write my lips: Interpretations of the relationship between German Sign Language and German across scales of SignWriting practice. *Signs and Society* 1.2: 243–272. DOI: 10.1086/672321

Holsanova, Jana, Nils Holmberg & Kenneth Holmqvist (2009): Reading information graphics: the role of spatial contiguity and dual attentional guidance. *Applied Cognitive Psychology* 23.9: 1215–1226. DOI: 10.1002/acp.1525

Hoosain, Rumjahn (1992): Psychological reality of the word in Chinese. In Chen Hsuan-Chih & Ovid. J. L. Tzeng (eds.), *Language processing in Chinese* (= Advances in Psychology, 90), 111–130. Amsterdam: Elsevier. DOI: 10.1016/S0166-4115(08)61889-0

Hopkins, Jason (2008): Choosing how to write sign language: a sociolinguistic perspective. *International Journal of the Sociology of Language* 192: 75–89. DOI: 10.1515/IJSL.2008.036

Houston, Stephen, John Baines & Jerrold Cooper (2003): Last writing: script obsolescence in Egypt, Mesopotamia, and Mesoamerica. *Comparative Studies in Society and History* 45.3: 430–479. DOI: 10.1017/S0010417503000227

Hsieh, Feng-fan (2015): Transcribing foreign names. In Rint Sybesma (ed.), *Encyclopedia of Chinese language and linguistics*. Leiden: Brill. DOI: 10.1163/2210-7363_ecll_COM_00000182

Hu, Zhuanglin (2015): The standardization of Chinese characters. *Chinese Semiotic Studies* 11.2: 123–133. DOI: 10.1515/css-2015-0007

Hurch, Bernhard (1988): Über Aspiration: Ein Kapitel aus der natürlichen Phonologie. Tübingen: Gunter Narr.

Hurch, Bernhard (2006): Natural Phonology. In Keith Brown (ed.), *Encyclopedia of language and linguistics*, 541–543. Amsterdam: Elsevier.

Hurch, Bernhard & Geoffrey Nathan (1996): Naturalness in phonology. *Sprachtypologie und Universalienforschung* 49.3: 231–245. DOI: 10.1524/stuf.1996.49.3.231

Ibbotson, Paul (2013): The scope of usage-based theory. *Frontiers in Psychology* 4: 255. DOI: 10.3389/fpsyg.2013.00255

Ibrahim, Raphiq (2013): Reading in Arabic: new evidence for the role of vowel signs. *Creative Education* 4.4: 248–253. DOI: 10.4236/ce.2013.44036

Inkelas, Sharon, Keith Johnson, Charles Lee, Emil Minas, George Mulcaire, Gek Yong Keng & Tomomi Yuasa (2013): Testing the learnability of writing systems. *Proceedings of the 39th Annual Meeting of the Berkeley Linguistics Society*, 75–89. Berkeley: Berkeley Linguistics Society. DOI: 10.3765/bls.v39i1.3871

Jacobs, Joachim (2005): *Spatien. Zum System der Getrennt- und Zusammenschreibung im heutigen Deutsch* (= Linguistik – Impulse & Tendenzen, 8). Berlin, Boston: De Gruyter. DOI: 10.1515/9783110919295

Jaffe, Alexandra M., Jannis Androutsopoulos, Mark Sebba & Sally Johnson (eds.) (2012): *Orthography as social action* (= Language and Social Processes, 3). Berlin, Boston: De Gruyter. DOI: 10.1515/9781614511038

James, Carl [1998] (2013): *Errors in language learning and use: exploring error analysis* (= Applied Linguistics and Language Study). London, New York: Routledge.

Jeffery, Lilian H. (1961): *The local scripts of archaic Greece. A study of the origin of the Greek alphabet and its development from the eighth to the fifth centuries B.C.* Oxford: OUP.

Jespersen, Bjørn & Chris Reintges (2008): Tractarian Sätze, Egyptian hieroglyphs, and the very idea of script as a picture. *The Philosophical Forum* 39.1: 1–19. DOI: 10.1111/j.1467-9191.2007.00279.x

Johnson, Sally (2002): On the origin of linguistic norms: orthography, ideology and the first constitutional challenge to the 1996 reform of German. *Language in Society* 31: 549–576. DOI: 10.1017/S0047404502314039

Johnson, Sally (2005): *Spelling trouble? Language, ideology and the reform of German orthography*. Bristol: Multilingual Matters.

Johnson, Rebecca L., Becky Bui & Lindsay L. Schmitt (2018): Are two spaces better than one? The effect of spacing following periods and commas during reading. *Attention, Perception, & Psychophysics* 80: 1504–1511. DOI: 10.3758/s13414-018-1527-6

Jones, Mari C. & Damien Mooney (2017): Creating orthographies for endangered languages. In Mari C. Jones & Damien Mooney (eds.), *Creating orthographies for endangered languages*, 1–35. Cambridge: CUP.

Joyce, Terry (2011): The significance of the morphographic principle for the classification of writing systems. *Written Language and Literacy* 14.1: 58–81. DOI: 10.1075/wll.14.1.04joy

Joyce, Terry & Susanne R. Borgwaldt (2011): Typology of writing systems: special issue introduction. *Written Language and Literacy* 14.1: 1–11. DOI: 10.1075/wll.14.1.01joy

Judson, Anna P. (2017): Review of *Variation within and among writing systems*, ed. by P. Cotticelli-Kurras & A. Rizza. *Bryn Mawr Classical Review* 2017.08.38. <http://bmcr.brynmawr.edu/2017/2017-08-38.html> (February 9th, 2019).

Kager, René (1999): *Optimality theory*. Cambridge: CUP.

Kanbe, Fumio (2013): On the generality of the topological theory of visual shape perception. *Perception* 42.8: 849–872. DOI: 10.1068/p7497

Kandel, Sonia, Carlos J. Álvarez & Nathalie Vallée (2006): Syllables as processing units in handwriting production. *Journal of Experimental Psychology: Human Perception and Performance* 32.1: 18–31. DOI: 10.1037/0096-1523.32.1.18

Kandel, Sonia, Lucie Hérault, Géraldine Grosjacques, Eric Lambert & Michel Fayol (2009): Orthographic vs. phonologic syllables in handwriting production. *Cognition* 110.3: 440–444. DOI: 10.1016/j.cognition.2008.12.001

Kao, Henry S. R., Gerard P. Van Galen & Rumjahn Hoosain (eds.) (1986): *Graphonomics: contemporary research in handwriting* (= Advances in Psychology, 37). Amsterdam: Elsevier Science.

Karan, Elke (2014): Standardization: what's the hurry? In Michael Cahill & Keren Rice (eds.), *Developing orthographies for unwritten languages*, 107–138. Dallas: SIL International.

Karg, Ina (2015): *Orthographie: Öffentlichkeit, Wissenschaft und Erwerb* (= Germanistische Arbeitshefte, 46). Berlin, Boston: De Gruyter. DOI: 10.1515/9783110366679

Kastl, Albert J. & Irvin L. Child (1968): Emotional meaning of four typographical variables. *Journal of Applied Psychology* 52.6,1: 440–446. DOI: 10.1037/h0026506

Katz, Leonard & Ram Frost (1992): The reading process is different for different orthographies: the orthographic depth hypothesis. In Ram Frost & Leonard Katz (eds.), *Orthography, phonology, morphology, and meaning*, 67–84. Amsterdam: Elsevier. DOI: 10.1016/S0166-4115(08)62789-2

Keage, Hannah A. D., Scott Coussens, Mark Kohler, Myra Thiessen & Owen F. Churches (2014): Investigating letter recognition in the brain by varying typeface: An event-related potential study. *Brain and Cognition* 88: 83–89. DOI: 10.1016/j.bandc.2014.05.001

Keller, Rudi (2014): *Sprachwandel: von der unsichtbaren Hand in der Sprache*, 4th edition (= UTB, 1567). Tübingen: Francke.

Kenner, Charmian (2004): *Becoming biliterate: young children learning different writing systems*. Stoke on Trent, UK, Sterling, US: Trentham Books.

Khateb, Asaid, Haitham Y. Taha, Inas Elias & Raphiq Ibrahim (2013): The effect of the internal orthographic connectivity of written Arabic words on the process of the visual recognition: a comparison between skilled and dyslexic readers. *Writing Systems Research* 5.2: 214–233. DOI: 10.1080/17586801.2013.834244

Kilani-Schoch, Marianne (2001): Naturalness or preference: a note on terminology in Natural Morphology. In Chris Schaner-Wolles, John Rennison & Friedrich Neubarth (eds.), *Naturally! Linguistic studies in honour of Wolfgang Ulrich Dressler presented on the occasion of his 60th birthday* (= Linguistica, 19), 233–240. Torino: Rosenberg & Selier.

Kim, Sung-Do (2011): Iconicity of Korean writing: a media semiotic approach. In Antonio Loprieno, Carsten K. Salis & Birgit Mersmann (eds*.), Bild, Macht, Schrift: Schriftkulturen in bildkritischer Perspektive*, 171–199. Weilerswist: Velbrück Wissenschaft.

Kirby, Simon, Hannah Cornish & Kenny Smith (2008): Cumulative cultural evolution in the laboratory: an experimental approach to the origins of structure in human language. *PNAS* 105.31: 10681–10686. DOI: 10.1073/pnas.0707835105

Kirchhoff, Frank (2016): Interpunktion und Intonation. In Ulrike Domahs & Beatrice Primus (eds.), *Handbuch Laut, Gebärde, Buchstabe* (= Handbooks of Linguistic Knowledge, 2), 398–417. Berlin, Boston: De Gruyter. DOI: 10.1515/9783110295993-022

Kirchhoff, Frank (2017): *Von der Virgel zum Komma: Die Entwicklung der Interpunktion im Deutschen* (= Germanistische Bibliothek, 61). Heidelberg: Winter.

Kirchhoff, Frank & Beatrice Primus (2016): Punctuation. In Vivian Cook & Des Ryan (eds.), *The Routledge handbook of the English writing system* (= Routledge Handbooks in Linguistics), 95–111. London: Routledge.

Klinkenberg, Jean-Marie & Stéphane Polis (2018): On scripturology. *Signata: Annals of Semiotics* 9: 57–102.

Koch, Peter & Wulf Oesterreicher (1985): Sprache der Nähe – Sprache der Distanz: Mündlichkeit und Schriftlichkeit im Spannungsfeld von Sprachtheorie und Sprachgeschichte. *Romanistisches Jahrbuch* 36: 15–43.

Koch, Peter & Wulf Oesterreicher (1994): Schriftlichkeit und Sprache. In Hartmut Günther & Otto Ludwig (eds.), *Schrift und Schriftlichkeit/Writing and its use* (= Handbooks of Linguistics and Communication Science, 10.1), 587–604. Berlin, Boston: De Gruyter. DOI: 10.1515/9783110111293.1.2.587

Koch, Peter & Wulf Oesterreicher (2012): Language of Immediacy – Language of Distance: Orality and Literacy from the Perspective of Language Theory and Linguistic History. In Claudia Lange, Beatrix Weber & Göran Wolf (eds.), *Communicative spaces: Variation, contact, and change – Papers in honour of Ursula Schaefer*, 441–473. Frankfurt a. M.: Peter Lang. DOI: 10.15496/publikation-20415

Kohrt, Manfred (1985): *Problemgeschichte des Graphembegriffs und des frühen Phonembegriffs* (= Reihe Germanistische Linguistik; 61). Tübingen: Niemeyer. DOI: 10.1515/9783111371207

Kohrt, Manfred (1986): The term 'grapheme' in the history and theory of linguistics. In Gerhard Augst (ed.), *New trends in graphemics and orthography*, 80–96. Berlin, Boston: De Gruyter. DOI: 10.1515/9783110867329.80

Kohrt, Manfred (1987): *Theoretische Aspekte der deutschen Orthographie* (= Reihe Germanistische Linguistik, 70). Berlin, Boston: De Gruyter. DOI: 10.1515/9783111371580

Kohrt, Manfred (1990): Die 'doppelte Kodifikation' der deutschen Orthographie. In Christian Stetter (ed.), *Zu einer Theorie der Orthographie: interdisziplinäre Aspekte gegenwärtiger Schrift- und Orthographieforschung* (= Reihe Germanistische Linguistik, 99), 104–144. Berlin, Boston: De Gruyter. DOI: 10.1515/9783111372280.104

Kolers, Paul A. (1983): Polarization of reading performance. In Florian Coulmas & Konrad Ehlich (eds.), *Writing in focus* (= Trends in Linguistics, 24), 371–391. Berlin, Boston: De Gruyter. DOI: 10.1515/9783110822830.371

König, Anne R. (2004): *Lesbarkeit als Leitprinzip der Buchtypographie. Eine Untersuchung zum Forschungsstand und zur historischen Entwicklung des Konzeptes ‚Lesbarkeit'* (= Studien der Erlanger Buchwissenschaft, 7). Erlangen: Universität Erlangen-Nürnberg. <http://alles-buch.uni-erlangen.de/Koenig.pdf> (February 6th, 2018).

Krämer, Sibylle (2001): *Sprache, Sprechakt, Kommunikation. Sprachtheoretische Positionen des 20. Jahrhunderts*. Frankfurt a. M.: Suhrkamp.

Krämer, Sybille (2003): ›Schriftbildlichkeit‹ oder: Über eine (fast) vergessene Dimension der Schrift. In Sybille Krämer & Horst Bredekamp (eds.), *Bild, Schrift, Zahl* (= Kulturtechnik), 157–176. München: Wilhelm Fink.

Ladd, D. Robert (2014): *Simultaneous structure in phonology*. Oxford: OUP.

Ladefoged, Peter (1997): Linguistic phonetic descriptions. In William J. Hardcastle & John Laver (eds.), *The handbook of phonetic sciences*, 589–618. Oxford: Wiley-Blackwell.

Lanthier, Sophie N., Evan F. Risko, Jennifer A. Stolz & Derek Besner (2009): Not all visual features are created equal: early processing in letter and word recognition. *Psychonomic Bulletin & Review* 16.1: 67–73. DOI: 10.3758/PBR.16.1.67

Laver, John (2017): Linguistic phonetics: the sounds of languages. In Mark Aronoff & Janie Rees-Miller (eds.), *The handbook of linguistics*, 2nd edition, 150–184. Oxford: Wiley-Blackwell. DOI: 10.1002/9781119072256.ch9

Lavine, Linda O. (1977): Differentiation of letterlike forms in prereading children. *Developmental Psychology* 13.2: 89–94. DOI: 10.1037/0012-1649.13.2.89

Law, Nancy, W.W. Ki, A.L.S. Chung, P.Y. Ko & H.C. Lam (1998): Children's stroke sequence errors in writing Chinese characters. *Reading and writing* 10.3–5: 267–292. DOI: 10.1023/A:1008091730338

Lee, Sang-Oak (2009): The Korean alphabet: An optimal featural system with graphical ingenuity. *Written Language and Literacy* 12.2: 202–212. DOI: 10.1075/wll.12.2.05lee

Lehtonen, Annukka & Peter Bryant (2005): Doublet challenge: form comes before function in children's understanding of their orthography. *Developmental Science* 8.3: 211–217. DOI: 10.1111/j.1467-7687.2005.00409.x

Levin, Iris & Ofra Korat (1993): Sensitivity to phonological, morphological and semantic cues in early reading and writing in Hebrew. *Merrill-Palmer Quarterly* 39.2: 213–232.

Levy, Betty A., Zhiyu Gong, Sandra Hessels, Mary A. Evans & Debra Jared (2006): Understanding print: early reading development and the contributions of home literacy experiences. *Journal of Experimental Child Psychology* 93.1: 63–93. DOI: 10.1016/j.jecp.2005.07.003

Li, David C. S. (2018): Writing Chinese: a challenge for Cantonese-L1 and South Asian Hongkongers. In Constanze Weth & Kasper Juffermans (eds.), *The tyranny of writing: ideologies of the written word*, 149–163. London, New York: Bloomsbury.

Li, Julia X. & Karin H. James (2016): Handwriting generates variable visual input to facilitate symbol learning. *Journal of Experimental Psychology: General* 145.3: 298–313. DOI: 10.1037/xge0000134

Lieber, Rochelle (2014): Natürliche Morphologie. In Susan Olsen & Peter O. Müller (eds.), *Wortbildung/Word Formation* (= Dictionaries of Linguistics and Communication Science, 2). Berlin, Boston: De Gruyter. DOI: 10.1515/wsk.2.0.naturlichemorphologie

Lindqvist, Christer (2001): *Skandinavische Schriftsysteme im Vergleich* (= Linguistische Arbeiten, 430). Berlin, Boston: De Gruyter. DOI: 10.1515/9783110927085

Linell, Per (1982): *The written language bias in linguistics*. Linköping: University of Linköping.

Liu, Pingping & Qin Lu (2018): The effects of spaces on word segmentation in Chinese reading: evidence from eye movements. *Journal of Research in Reading* 41.2: 329–349. DOI: 10.1111/1467-9817.12106

Lockwood, David G. (2001): Phoneme and grapheme: how parallel can they be? *LACUS Forum* 27: 307–316.

Logan, Robert K. (2004): *The Alphabet Effect: a media ecology understanding of the making of Western civilization*. New York: Hampton Press.

Longcamp, Marieke, Marie-Thérèse Zerbato-Poudou & Jean-Luc Velay (2005): The influence of writing practice on letter recognition in preschool children: a comparison between handwriting and typing. *Acta Psychologica* 119.1: 67–79. DOI: 10.1016/j.actpsy.2004.10.019

Longcamp, Marieke, Yevhen Hlushchuk & Riita Hari (2011): What differs in visual recognition of handwritten vs. printed letters? An fMRI study. *Human Brain Mapping* 32.8: 1250–1259. DOI: 10.1002/hbm.21105

Ludwig, Otto (2007): Skripte. Konturen einer Konzeption. *Zeitschrift für germanistische Linguistik* 35.3: 376–396. DOI: 10.1515/zgl.2007.025

Lund, Ole (1999): *Knowledge construction in typography: the case of legibility research and the legibility of sans serif typefaces*. PhD dissertation, University of Reading.

Lüpke, Friederike (2011): Orthography development. In Peter Austoin & Julia Sallabank (eds.), *Handbook of endangered languages*, 312–336. Cambridge: CUP.

Luschützky, Hans C. (1991): *Twenty years of naturalism: a bibliography* (= Wiener Linguistische Gazette, Beiheft 10). Wien: Institut für Sprachwissenschaft der Universität Wien.

Ma, Guojie (2017): Does interword spacing influence lexical processing in Chinese reading? *Visual Cognition* 25.7–8: 815–824. DOI: 10.1080/13506285.2017.1338322

Maas, Utz (1992): *Grundzüge der deutschen Orthographie* (= Reihe Germanistische Linguistik, 120). Berlin, Boston: De Gruyter. DOI: 10.1515/9783111376974

Maas, Utz (2015): Vom Sinn der Rechtschreibung. *Networx* 68. <http://www.mediensprache.net/networx/networx-68.pdf> (December 29th, 2017).

Mair, Victor (2011): Polysyllabic characters in Chinese writing. *Language Log* <http://languagelog.ldc.upenn.edu/nll/?p=3330> (September 22nd, 2020).

Majidi, Mohammad-Reza (1996): *Einführung in die arabisch-persische Schrift*. Hamburg: Helmut Buske.

Mallon, Jean (1952): *Paléographie romaine* (= Scripturae, Monumenta et Studia, 3). Madrid: Instituto Antonio de Nebrija de Filologia.

Mangen, Anne (2016): What hands may tell us about reading and writing. *Educational Theory* 66.4: 457–477. DOI: 10.1111/edth.12183

Maroun, Maryse & J. Richard Hanley (2017): Diacritics improve comprehension of the Arabic script by providing access to the meanings of heterophonic homographs. *Reading and Writing* 30.2: 319–335. DOI: 10.1007/s11145-016-9677-1

Martinet, André (1949): La double articulation linguistique. *Travaux du Cercle Linguistique de Copenhague* 5: 30–37.

Massaro, Dom (2011): Reflections on the syllable as the perceptual unit in speech perception. In David Poeppel & Greg Hickok (eds.), *Talking Brains*. <http://www.talkingbrains.org/2011/02/reflections-on-syllable-as-perceptual.html> (September 22nd, 2020).

Mattingly, Ignatius G. (1985): Did orthographies evolve? *Remedial and Special Education* 6.6: 18–23. DOI: 10.1177/074193258500600605

Mayer, Richard E. (2005): Principles for managing essential processing in multimedia learning: coherence, signalling, redundancy, spatial contiguity and temporal contiguity principles. In Richard E. Mayer (ed.), *Cambridge handbook of multimedia learning*, 183–200. Cambridge: CUP.

Mayerthaler, Willi (1977): *Studien zur theoretischen und zur französischen Morphologie: Reduplikation, Echowörter, morphologische Natürlichkeit, Haplologie, Produktivität, Regeltelescoping, paradigmatischer Ausgleich* (= Linguistische Arbeiten, 40). Berlin, Boston: De Gruyter. DOI: 10.1515/9783111592503

Mayerthaler, Willi (1987): System-independent morphological naturalness. In Wolfgang U. Dressler, Willi Mayerthaler, Oswald Panagl & Wolfgang U. Wurzel, *Leitmotifs in Natural Morphology* (= Studies in Language Companion Series, 10), 25–58. Amsterdam, Philadelphia: John Benjamins. DOI: 10.1075/slcs.10.13may

McCardle, Peggy, Brett Miller, Jun Ren Lee & Ovid J. L. Tzeng (2011): *Dyslexia across languages: orthography and the brain-gene-behavior link*. Baltimore: Paul H. Brookes Publishing.

McCawley, James D. (1994): Some graphotactic constraints. In W. C. Watt (ed.), *Writing systems and cognition* (= Neuropsychology and Cognition, 6), 115–127. Dordrecht: Springer. DOI: 10.1007/978-94-015-8285-8_7

McClelland, James L. & David E. Rumelhart (1981): An Interactive Activation Model of context effects in letter perception: part 1. An account of basic findings. *Psychological Review* 88.5: 375–407. DOI: 10.1037/0033-295X.88.5.375

McIntosh, Robert D., Keira Hillary, Ailbhe Brennan & Magdalena Lechowicz (2018): Developmental mirror-writing is paralleled by orientation recognition errors. *Laterality: Asymmetries of Body, Brain and Cognition* 23.6: 664–683. DOI: 10.1080/1357650X.2018.1445748

Medwell, Jane & David Wray (2008): Handwriting – a forgotten language skill? *Language and Education* 22.1: 34–47. DOI: 10.2167/le722.0

Meier, Thomas, Michael R. Ott & Rebecca Sauer (2015): *Materiale Textkulturen: Konzepte – Materialien – Praktiken* (= Materiale Textkulturen, 1). Berlin, Boston: De Gruyter. DOI: 10.1515/9783110371291

Meisenburg, Trudel (1996): *Romanische Schriftsysteme im Vergleich: eine diachrone Studie* (= ScriptOralia, 82). Tübingen: Gunter Narr.

Meletis, Dimitrios (2015): *Graphetik: Form und Materialität von Schrift*. Glückstadt: Verlag Werner Hülsbusch.

Meletis, Dimitrios (2017): Review of *Handbuch Laut, Gebärde, Buchstabe*, ed. by U. Domahs & B. Primus. *Zeitschrift für Rezensionen zur germanistischen Sprachwissenschaft* 9.1–2, 109–115. DOI: 10.1515/zrs-2017-0019

Meletis, Dimitrios (2018): What is natural in writing? Prolegomena to a Natural Grapholinguistics. *Written Language and Literacy* 21.1: 52–88. DOI: 10.1075/wll.00010.mel

Meletis, Dimitrios (2019): The grapheme as a universal basic unit of writing. *Writing Systems Research* 11.1: 26–49. DOI: 10.1080/17586801.2019.1697412

Meletis, Dimitrios (2020a): Warum hassen alle Comic Sans? Metapragmatische Onlinediskurse zu einer typographischen Hassliebe. In Jannis Androutsopoulos & Florian Busch (eds.), *Register des Graphischen: Variation, Praktiken, Reflexion*, 235–267. Berlin, Boston: De Gruyter. DOI: 10.1515/9783110673241-010

Meletis, Dimitrios (2020b): Types of allography. *Open Linguistics* 6: 249–266. DOI: 10.1515/opli-2020-0006

Meletis, Dimitrios (accepted a): Reintroducing graphetics: The study of the materiality of writing. *Scripta*.

Meletis, Dimitrios (accepted b): Universality and diversity in writing systems. *LACUS Forum* 46.

Mesch, Birgit & Christina Noack (2016): System, Norm und Gebrauch – drei Seiten derselben

Medaille? In Birgit Mesch & Christina Noack (eds.), *System, Norm und Gebrauch – drei Seiten derselben Medaille? Orthographische Kompetenz und Performanz zwischen System, Norm und Empirie*, 1–8. Baltmannsweiler: Schneider Hohengehren.

Michel, Lothar (1996): Forensische Handschriftenuntersuchung. In Hartmut Günther & Otto Ludwig (eds.), *Schrift und Schriftlichkeit/Writing and its use* (= Handbooks of Linguistics and Communication Science, 10.2), 1036–1048. Berlin, Boston: De Gruyter. DOI: 10.1515/9783110147445.2.7.1036

Miestamo, Matti (2008): Grammatical complexity in a cross-linguistic perspective. In Matti Miestamo, Kaius Sinnemäki & Fred Karlsson (eds.), *Language complexity: typology, contact, change*, 23–42. Amsterdam, Philadelphia: John Benjamins. DOI: 10.1075/slcs.94.04mie

Milroy, Lesley (1982): Language and group identity. *Journal of Multilingual and Multicultural Development* 3.3: 207–216. DOI: 10.1080/01434632.1982.9994085

Miret, Fernando Sánchez (2009): Fundamentos de la Lingüística Natural. *Revista Portuguesa de Humanidades | Edstudos Linguísticos* 13.1: 179–190.

Miyamoto, Tadao (2007). The evolution of writing systems: against the Gelbian hypothesis. In Akito Sakurai, Kôiti Hasida & Katsumi Nitta (eds.), *New frontiers in artificial intelligence* (= Lecture Notes in Artificial Intelligence, 3609), 345–356. Berlin: Springer. DOI: 10.1007/978-3-540-71009-7_31

Morin, Olivier (2018): Spontaneous emergence of legibility in writing systems: the case of orientation anisotropy. *Cognitive Science* 42.2: 664–677. DOI: 10.1111/cogs.12550

Morris, Robin K., Keith Rayner & Alexander Pollatsek (1990): Eye movement guidance in reading. *Journal of Experimental Psychology: Human Perception and Performance* 16.2: 268–281. DOI: 10.1037/0096-1523.16.2.268

Moseley, Christopher (2014): *Writing systems* (= Critical Concepts in Linguistics). London: Routledge.

Mugdan, Joachim (1990): On the history of linguistic terminology. In Hans-Josef Niederehe & E. F. K. Koerner (eds.), *History and historiography of linguistics: proceedings of the fourth International Conference on the History of the Language Sciences (ICHoLS IV), Trier, 24–28 August 1987,* vol. 1: *Antiquity–17th century* (= Studies in the History of the Language Sciences, 51.1), 49–62. Amsterdam: John Benjamins. DOI: 10.1075/sihols.51.1.09mug

Mulder, Jan W. F. (1994): Written and spoken languages as separate semiotic systems. *Semiotica* 101.1–2: 41–72.

Mullaney, Thomas S. (2017): *The Chinese typewriter: a history*. Cambridge: MIT Press.

Munske, Horst H. (1994): Ist eine „natürliche“ Graphematik möglich? In Otmar Werner (ed.), *Probleme der Graphie* (= Script Oralia, 57), 9–24. Tübingen: Gunter Narr.

Myers, James (2019). *The grammar of Chinese characters: productive knowledge of formal patterns in an orthographic system*. London: Routledge.

Nag, Sonali, Margaret Snowling, Philip Quinlan & Charles Hume (2014): Child and symbol factors in learning to read a visually complex writing system. *Scientific Studies of Reading* 18.5: 309–324. DOI: 10.1080/10888438.2014.892489

Nathan, Geoffrey (1999): What functionalists can learn from formalists in phonology. In Michael Darnell, Edith A. Moravcsik, Michael Noonan, Frederick J. Newmeyer & Kathleen Wheatley (eds.), *Functionalism and formalism in linguistics,* vol. 1: *general papers* (= Studies in Language Companion Series, 41), 305–327. DOI: 10.1075/slcs.41.15nat

Nathan, Geoffrey (2007): Where is the Natural Phonology phoneme in 2007? In Jürgen Trouvain & William J. Barry (eds.), *Proceedings of the 16th International Congress of Phonetic Sciences*, 93–98. Saarbrücken: University of Saarbrücken.

Naumann, Carl L. (1989): *Gesprochenes Deutsch und Orthographie. Linguistische und didaktische Studien zur Rolle der gesprochenen Sprache im System und Erwerb der Rechtschreibung* (= Theorie und Vermittlung der Sprache, 8). Frankfurt a. M.: Peter Lang.

Naumann, Carl L. (1990): Nochmals zu den Prinzipien der Orthographie. In Christian Stetter (ed.), *Zu einer Theorie der Orthographie: interdisziplinäre Aspekte gegenwärtiger Schrift- und Orthographieforschung* (= Reihe Germanistische Linguistik, 99), 145–162. Boston, Berlin: De Gruyter. DOI: 10.1515/9783111372280.145

Neef, Martin (2005): *Die Graphematik des Deutschen* (= Linguistische Arbeiten, 500). Berlin, Boston: De Gruyter. DOI: 10.1515/9783110914856

Neef, Martin (2012): Graphematics as part of a modular theory of phonographic writing systems. *Writing Systems Research* 4.2: 214–228. DOI: 10.1080/17586801.2012.706658

Neef, Martin (2015): Writing systems as modular objects: proposals for theory design in grapholinguistics. *Open Linguistics* 1: 708–721. DOI: 10.1515/opli-2015-0026

Neef, Martin (2019): Rezension von Ulrike Domahs/Beatrice Primus (Hg.) (2016) Handbuch Laut, Gebärde, Buchstabe. Berlin/New York: De Gruyter. XX, 516 S. (Handbücher Sprachwissen. 2). *Zeitschrift für Dialektologie und Linguistik* 86.1: 101–106.

Neef, Martin & Miriam Balestra (2011): Measuring graphematic transparency – German and Italian compared. *Written Language and Literacy* 14.1: 109–142. DOI: 10.1075/wll.14.1.06nee

Neef, Martin, Said Sahel & Rüdiger Weingarten (eds.) (2012ff.): *Schriftlinguistik/Grapholinguistics* (= Dictionaries of Linguistics and Communication Science, 5). Berlin, Boston: De Gruyter.

Nerius, Dieter (2007): Merkmale der Orthographie. In Dieter Nerius (et al.) (eds.), *Deutsche Orthographie*, 4th edition, 34–40. Hildesheim: Olms.

Nerius, Dieter (2012): Zur Geschichte der Schriftlinguistik in der Germanistik der DDR. In Jan Cölln & Franz-Josef Holznagel (eds.), *Positionen der Germanistik in der DDR: Personen – Forschungsfelder – Organisationsformen*, 387 – 397. Berlin, Boston: De Gruyter. DOI: 10.1515/9783110223842.387

Nerius, Dieter (2020): Kodifikation der Schreibung. In Thomas Nier, Jörg Kilian & Jürgen Schiewe (eds.), *Handbuch Sprachkritik*, 368–375. Stuttgart: Metzler.

Nerius, Dieter & Gerhard Augst (eds.) (1988): *Probleme der geschriebenen Sprache. Beiträge zur Schriftlinguistik auf dem XIV. internationalen Linguistenkongreß 1987 in Berlin* (= Linguistische Studien, A, 173). Berlin: Akademie der Wissenschaften der DDR.

Nöth, Winfried (1990): *Handbook of semiotics* (= Advances in Semiotics). Indiana: Indiana University Press.

Nöth, Winfried (2001): Semiotic foundations of iconicity in language and literacy. In Olga Fischer & Max Nänny (eds.), *The motivated sign* (= Iconicity in Language and Literature, 2), 17–28. Amsterdam, Philadelphia: John Benjamins. DOI: 10.1075/ill.2.05not

Nottbusch, Guido, Angela Grimm, Rüdiger Weingarten & Udo Will (2005): Syllabic structures in typing: evidence from deaf writers. *Reading and Writing* 18.6: 497–526. DOI: 10.1007/s11145-005-3178-y

Noyes, Jan (1983): The QWERTY keyboard: a review. *International Journal of Man-Machine Studies* 18.3: 265–281. DOI: 10.1016/S0020-7373(83)80010-8

Nunberg, Geoffrey, Ted Briscoe & Rodney Huddleston (2002): Punctuation. In Rodney Huddleston & Geoffrey Pullum (eds.), *The Cambridge grammar of the English language*, 1723–1764. Cambridge: CUP.

Olson, David R. (1996): *The world on paper: the conceptual and cognitive implications of writing and reading*. Cambridge: CUP.

Ong, Walter J. (2012): *Orality and literacy: the technologizing of the word*, 30th anniversary edition, with additional chapters by John Hartley. London, New York: Routledge.

O'Regan, J. Kevin (1990): Eye movements and reading. In Eileen Kowler (ed.), *Eye movements and their role in visual and cognitive processes* (= Reviews of Oculomotor Research, 4), 395–453. Amsterdam: Elsevier.

O'Shea, Michael (2012): *Das Gehirn* (transl. by Manfred Weltecke), 2nd edition. Stuttgart: Reclam.

Otake, Shoko, Rebecca Treiman & Li Yin (2017): Differentiation of writing and drawing by U.S. two- to five-year-olds. *Cognitive Development* 43: 119–128. DOI: 10.1016/j.cogdev.2017.03.004

Pabst-Weinschenk, Marita (2016): Orthoepie. In Jutta Rymarczyk & Jörg Kilian (eds.), *Sprachdidaktik: Erstsprache, Zweitsprache, Fremdsprache/Language learning: first language, second language, foreign language* (= Dictionaries of Linguistics and Communication Science, 16). Berlin, Boston: De Gruyter. DOI: 10.1515/wsk.16.0.orthoepie

Palmer, Tony (2015): *The practice of Chinese typography*, vol. 1: *typefaces*. Melbourne: ACP.

Parkinson, Jim, Benjamin J. Dyson & Beena Khurana (2010): Line by line: the ERP correlates of stroke order priming in letters. *Experimental Brain Research* 201.3: 575–586. DOI: 10.1007/s00221-009-2077-x

Parush, Shula, Vered Pindak, Jeri Hahn-Markowitz & Tal Mazor-Karsenty (1998): Does fatigue influence children's handwriting performance? *Work* 11.3: 307–313. DOI: 10.3233/WOR-1998-11307

Patel, Purushottam G. (2010): *The Brahmi writing system: Cross-fertilizing epigraphy, archaeology and linguistics*. New Delhi: Black & White.

Paul-Mengelberg, Maria (1996): Graphologie. In Hartmut Günther & Otto Ludwig (eds.), *Schrift und Schriftlichkeit/Writing and its use* (= Handbooks of Linguistics and Communication Science, 10.2), 1049–1056. Berlin, Boston: De Gruyter. DOI: 10.1515/9783110147445.2.7.10 49

Paz-Villagrán, Vietminh, Jérémy Danna & Jean-Luc Velay (2014): Lifts and stops in proficient and dysgraphic writing. *Human Movement Science* 33: 381–394. DOI: 10.1016/j.humov.2013.11.005

Pegado, Filipe, Kimihiro Nakamura, Laurent Cohen & Stanislas Dehaene (2011): Breaking the symmetry: Mirror discrimination for single letters but not for pictures in the Visual Word Form Area. *NeuroImage* 55.2: 742–749. DOI: 10.1016/j.neuroimage.2010.11.043

Pelli, Denis G., Catherine W. Burns, Bart Farell & Deborah C. Moore-Page (2006): Feature detection and letter identification. *Vision Research* 46.28: 4646–4674. DOI: 10.1016/j.visres.2006.04.023

Perea, Manuel & Manuel Carreiras (2006): Do transposed-letter effects occur across lexeme boundaries? *Psychonomic Bulletin & Review* 13.3: 418–422. DOI: 10.3758/BF03193863

Perea, Manuel & Joana Acha (2009): Space information is important for reading. *Vision Research* 49.15: 1994–2000. DOI: 10.1016/j.visres.2009.05.009

Perea, Manuel, Reem Abu Mallouh & Manuel Carreiras (2010): The search for an input-coding scheme: transposed-letter priming in Arabic. *Psychonomic Bulletin & Review* 17.3: 375–380. DOI: 10.3758/PBR.17.3.375

Perea, Manuel, Chie Nakatani & Cees van Leeuwen (2011): Transposition effects in reading Japanese Kana: are they orthographic in nature? *Memory & Cognition* 39.4: 700–707. DOI: 10.3758/s13421-010-0052-1

Perfetti, Charles A. & Lindsay N. Harris (2013): Universal reading processes are modulated by language and writing system. *Language Learning and Development* 9.4: 296–316. DOI: 10.1080/15475441.2013.813828

Petitto, Laura A., Clifton Langdon, Adam Stone, Diana Andriola, Geo Kartheiser & Casey Cochran (2016): Visual sign phonology: insights into human reading and language from a natural soundless phonology. *WIREs Cognitive Science* 7: 366–381. DOI: 10.1002/wcs.1404

Pettersson, John S. (1996): *Grammatological studies: writing and its relation to speech* (= Reports from Uppsala University Linguistics, 29). Uppsala: University of Uppsala.

Pike, Kenneth L. [1954] (1967): *Language in relation to a unified theory of the structure of human behavior* (= Janua Linguarum, Series Maior, 24). Berlin, Boston: De Gruyter. DOI: 10.1515/9783111657158

Pinheiro, Ângela M. V. (1995): Reading and spelling development in Brazilian Portuguese. *Reading and Writing* 7.1: 111–138. DOI: 10.1007/BF01026950

Platt, Penny (1977): Grapho-Linguistics: children's drawings in relation to reading and writing skills. *The Reading Teacher* 31.3: 262–268.

Pollo, Tatiana C., Brett Kessler & Rebecca Treiman (2009): Statistical patterns in children's early writing. *Journal of Experimental Child Psychology* 104.4: 410–426. DOI: 10.1016/j.jecp.2009.07.003

Pontecorvo, Clotilde (1985): Figure, parole, numeri: un problema di simbolizzazione. *Età Evolutiva* 22: 5–33.

Primus, Beatrice (2003): Zum Silbenbegriff in der Schrift-, Laut- und Gebärdensprache – Versuch einer mediumübergreifenden Fundierung. *Zeitschrift für Sprachwissenschaft* 23.1: 3–55. DOI: 10.1515/zfsw.2003.22.1.3

Primus, Beatrice (2004): A featural analysis of the Modern Roman Alphabet. *Written Language and Literacy* 7.2: 235–274. DOI: 10.1075/wll.7.2.06pri

Primus, Beatrice (2006): Buchstabenkomponenten und ihre Grammatik. In Ursula Bredel & Hartmut Günther (eds.), *Orthographietheorie und Rechtschreibunterricht* (= Linguistische Arbeiten, 509), 5–43. Berlin, Boston: De Gruyter. DOI: 10.1515/9783110921199.5

Primus, Beatrice (2010): Strukturelle Grundlagen des deutschen Schriftsystems. In Ursula Bredel & Astrid Müller & Gabriele Hinney (eds.), *Schriftsystem und Schrifterwerb: linguistisch – didaktisch – empirisch* (= Reihe Germanistische Linguistik, 289), 9–46. Berlin, Boston: De Gruyter. DOI: 10.1515/9783110232257.9

Primus, Beatrice & Lutz Wagner (2013): Buchstabenkomposition. In Ryôsuke Ôhashi & Martin Roussel (eds.), *Buchstaben der Welt – Welt der Buchstaben* (= Morphomata, 12), 33–58. München: Wilhelm Fink.

Punsongserm, Rachapoom, Shoji Sunaga & Hisayasu Ihara (2017): Thai typefaces (part 1): assumption on visibility and legibility problems. *Archives of Design Research* 30.1: 5–22. DOI: 10.15187/adr.2017.02.30.1.5

Qiu, Xigui (2000): *Chinese writing* (transl. by Gilbert L. Mattos, & Jerry Norman). Berkeley, CA: The Society for the Study of Early China.

Qian, Yi, Yao-Wu Song, Jing Zhao & Hong-Yan Bi (2015): The developmental trend of orthographic awareness in Chinese preschoolers. *Reading and Writing* 28.4: 571–586. DOI: 10.1007/s11145-014-9538-8

Rahbari, Noriyeh, Monique Sénéchal & Narges Arab-Moghaddam (2007): The role of orthographic and phonological processing skills in the reading and spelling of monolingual Persian children. *Reading and Writing* 20.5: 511–533. DOI: 10.1007/s11145-006-9042-x

Rahnenführer, Ilse (1989): Nochmals zum Status der Prinzipien in der Orthographie. In Peter Eisenberg & Hartmut Günther (eds.), *Schriftsystem und Orthographie* (= Reihe Germanistische Linguistik, 97), 283–296. Berlin, Boston: De Gruyter. DOI: 10.1515/9783111372266.283

Rayner, Keith (1979): Eye movements in reading: eye guidance and integration. In Paul A. Kolers, Merald E. Wrolstad & Herman Bouma (eds.), *Processing of visible language*, vol. 1 (= NATOCS, 13), 61–75. New York, London: Plenum Press. DOI: 10.1007/978-1-4684-0994-9_4

Rayner, Keith (1998): Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin* 124.3: 372–422. DOI: 10.1037/0033-2909.124.3.372

Rayner, Keith, Arnold D. Well, Alexander Pollatsek (1980): Asymmetry of the effective visual field in reading. *Perception & Psychophysics* 27.6: 537–544. DOI: 10.3758/BF03198682

Rayner, Keith, Martin H. Fischer & Alexander Pollatsek (1998): Unspaced text interferes with both word identification and eye movement control. *Vision Research* 38.8: 1129–1144. DOI: 10.1016/S0042-6989(97)00274-5

Rechtschreibrat (2018): *Bericht und Vorschläge der AG 'Geschlechtergerechte Schreibung' zur Sitzung des Rats für deutsche Rechtschreibung am 16.11.2018. Revidierte Fassung aufgrund des Beschlusses des Rats vom 16.11.2018*. <http://www.rechtschreibrat.com/DOX/rfdr_2018-11-28_anlage_3_bericht_ag_geschlechterger_schreibung.pdf> (September 22nd, 2020).

Rehg, Kenneth L. (2004): Linguistics, literacy, and the law of unintended consequences. *Oceanic Linguistics* 43.2: 498–518.

Reicher, Gerald M. (1969): Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology* 81.2: 274–280. DOI: 10.1037/h0027768

Reilly, Ronan, Ralph Radach, Daniel Corbic & Sudaporn Luksaneeyanawin (2005): Comparing reading in Thai and English – the role of spatial word unit segmentation in distributed processing and eye movement control. Presentation at the *13th European conference on eye movements* at the University of Bern, Switzerland. DOI: 10.16910/jemr.1.4.1

Reiner, Erica (2000): The Sumerian and Akkadian linguistic tradition. In Sylvain Auroux, E. F. K. Koerner, Hans-Josef Niederehe, Kees Versteegs (eds.), *History of the language sciences* (= Handbooks of Linguistics and Communication Science, 18.1), 1–5. Berlin, Boston: De Gruyter. DOI: 10.1515/9783110111033.1.1.1

Reißig, Tilo (2015): *Typographie und Grammatik: Untersuchung zum Verhältnis von Syntax und Raum* (= Stauffenburg Linguistik, 84). Tübingen: Stauffenburg.

Reißig, Tilo & Tobias Bernasconi (2015): Das Lesen typographischer Dispositive. In Iris Rautenberg & Tilo Reißig (eds.), *Lesen und Lesedidaktik aus linguistischer Perspektive* (= Forum Angewandte Linguistik, 53), 217–241. Frankfurt a. M.: Peter Lang.

Reitz, Jela. (1994): *Erworbene Schriftsprachstörungen: eine neurolinguistische Aufgabensammlung zur Erfassung schriftsprachlicher Leistungen*. Opladen, Wiesbaden: Westdeutscher Verlag.

Rey, Arnaud, Stéphane Dufau, Stéphanie Massol & Jonathan Grainger (2009): Testing computational models of letter perception with item-level event-related potentials. *Cognitive Neuropsychology* 26.1: 7–22. DOI: 10.1080/09541440802176300

Rezec, Oliver (2009): *Zur Struktur des deutschen Schriftsystems*. PhD Dissertation, Ludwig-Maximilians-Universität München.

Rezec, Oliver (2010): Der vermeintliche Zusammenhang zwischen Buchstabenformen und Lautwerten. Erwiderung auf einige Thesen von B. Primus. *Linguistische Berichte* 223: 343–366.

Rezec, Oliver (2013): Ein differenzierteres Strukturmodell des deutschen Schriftsystems. *Linguistische Berichte* 234: 227–254.

Rimzhim, Anurag, Leonard Katz & Carol A. Fowler (2014): Brāhmī-derived orthographies are typologically Āksharik but functionally predominantly alphabetic. *Writing Systems Research* 6.1: 41–53. DOI: 10.1080/17586801.2013.855618

Rizza, Alfredo (2018): Recurring problems in the application of the term/concept grapheme: reflecting again. Presentation at the *LSS München/Graphematik-Tagung des*

*Akademieprojekts RuneS „Wege zur Konfiguration der Zeichen-Phonem-Beziehung"* at LMU München, Munich, Germany, November 8th, 2018.

Robins, Sarah, Rebecca Treiman, Nicole Rosales & Shoko Otake (2012): Parent-child conversations about letters and pictures. *Reading and Writing* 25.8: 2039–2059. DOI: 10.1007/s11145-011-9344-5

Rogers, Henry (1995): Optimal orthographies. In Insup Taylor & David R. Olson (eds.), *Scripts and literacy: reading and learning to read alphabets, syllabaries and characters* (= Neuropsychology and Cognition, 7), 31–43. Dordrecht: Springer. DOI: 10.1007/978-94-011-1162-1_3

Rogers, Henry (2005): *Writing systems: a linguistic approach*. Oxford: Wiley-Blackwell.

Roldán, Manuel, Ana Marcet & Manuel Perea (2018): Is there a cost at encoding words with joined letters during visual word recognition? *Psicológica* 39.2: 279–291. DOI: 10.2478/psicolj-2018-0012

Rothlein, David & Brenda Rapp (2017): The role of allograph representations in font-invariant letter identification. *Journal of Experimental Psychology: Human Perception and Performance* 43.7: 1411–1429. DOI: 10.1037/xhp0000384

Ruf, Oliver (2014): *Wischen & Schreiben. Von Mediengesten zum digitalen Text*. Berlin: Kulturverlag Kadmos.

Rumelhart, David E. & Donald E. Norman (1982): Simulating a skilled typist: a study of skilled cognitive-motor performance. *Cognitive Science* 6.1: 1–36. DOI: 10.1207/s15516709cog0601_1

Rupareliya, Chintan, Syeda Naqvi & Seyedali Hejazi (2017): Alexia without agraphia: a rare entity. *Cureus* 9.6: e1304. DOI: 10.7759/cureus.1304

Ryding, Karin C. (2014): *Arabic: a linguistic introduction*. Cambridge: CUP.

Saenger, Paul (1991): The separation of words and the physiology of reading. In David R. Olson & Nancy Torrance (eds.), *Literacy and orality*, 198–214. Cambridge: CUP.

Saenger, Paul (1997): *Space between words: the origins of silent reading*. Stanford: Stanford University Press.

Saiegh-Haddad, Elinor & Roni Henkin-Roitfarb (2014): The structure of Arabic language and orthography. In Elinor Saiegh-Haddad & R. Malatesha Joshi (eds.), *Handbook of Arabic literacy: insights and perspectives* (= Literacy Studies, 9), 3–28. Dordrecht: Springer. DOI: 10.1007/978-94-017-8545-7_1

Saiegh-Haddad, Elinor & Bernard Spolsky (2014): Acquiring literacy in a diglossic context: problems and prospects. In Elinor Saiegh-Haddad & R. Malatesha Joshi (eds.), *Handbook of Arabic literacy: insights and perspectives* (= Literacy Studies, 9), 225–240. Dordrecht: Springer. DOI: 10.1007/978-94-017-8545-7_10

Saiegh-Haddad, Elinor & Rachel Schiff (2016): The impact of diglossia on voweled and unvoweled reading in Arabic: a developmental study from childhood to adolescence. *Scientific Studies of Reading* 20.4: 311–324. DOI: 10.1080/10888438.2016.1180526

Sainio, Miia, Jukka Hyönä, Kazuo Bingushi & Raymond Bertram (2007): The role of interword spacing in reading Japanese: an eye movement study. *Vision Research* 47.20: 2575–2584. DOI: 10.1016/j.visres.2007.05.017

Sakamoto, T. & Kiyoshi Makita (1973): Japan. In John Downing (ed.), *Comparative reading: cross-national studies of behavior in reading and writing*, 440–465. New York: Macmillan.

Salomon, Richard (2007): Typological observations on the Indic scripts and their relationship with other alphasyllburies [sic]. In Purushottam G. Patel, Pramod Pandey & Dilip Rajgor (eds.), *The Indic scripts: palaeographic and linguistic perspectives*, 25–43. New Delhi: D. K. Printworld.

Salomon, Richard (2012): Some principles and patterns of script change. In Stephen D. Houston (ed.), *The shape of script: how and why writing systems change*, 119–133. Santa Fe: School for Advanced Research Press.

Sampson, Geoffrey (1994): Chinese script and the diversity of writing systems. *Linguistics* 32: 117–132. DOI: 10.1515/ling.1994.32.1.117

Sampson, Geoffrey [1985] (2015): *Writing systems: a linguistic introduction*, 2nd edition. Bristol: Equinox.

Sampson, Geoffrey (2016): Typology and the study of writing systems. *Linguistic Typology* 20.3: 561–567. DOI: 10.1515/lingty-2016-0027

Sampson, Geoffrey (2018): From phonemic spelling to distinctive spelling. *Written Language and Literacy* 21.1: 3–25. DOI: 10.1075/wll.00008.sam

Sandler, Wendy (2012): The phonological organization of sign languages. *Language and Linguistics Compass* 6.3: 162–182. DOI: 10.1002/lnc3.326

Sanocki, Thomas (1988): Font regularity constraints on the process of letter recognition. *Journal of Experimental Psychology: Human Perception and Performance* 14.3: 472–480. DOI: 10.1037/0096-1523.14.3.472

Sanocki, Thomas (1991): Effects of early common features on form recognition. *Perception & Psychophysics* 50.5: 490–497. DOI: 10.3758/BF03205065

Sanocki, Thomas & Mary C. Dyson (2012): Letter processing and font information during reading: beyond distinctiveness, where vision meets design. *Attention, Perception, & Psychophysics* 74.1: 132–145. DOI: 10.3758/s13414-011-0220-9

Sapir, Edward (1933): La réalité psychologique des phonèmes. *Journal de psychologie normale et pathologique* 30: 247–265.

Sariti, Anthony W. (1967): *Chinese grapholinguistics*. Master's Thesis, Georgetown University.

Schaeder, Burkhard (1986): Das letzte Wort hat immer der Duden. Zu Geschichte, Inhalt und Funktion der Rechtschreib-Wörterbücher des Deutschen. In H. E. Wiegand (ed.), *Studien zur neuhochdeutschen Lexikographie,* VI. 1. Teilband (= Germanistische Linguistik 84–86/1986), 197–241. Hildesheim: Olms.

Schane, Sanford A. (1971): The phoneme revisited. *Language* 47.3: 503–521.

Schimmel-Fijalkowytsch, Nadine (2018): *Diskurse zur Normierung und Reform der deutschen Rechtschreibung. Eine Analyse von Diskursen zur Rechtschreibreform unter soziolinguistischer und textlinguistischer Perspektive* (= Studien zur deutschen Sprache, 75). Tübingen: Narr.

Schlieben-Lange, Brigitte (1996): Geschichte der Reflexion über Schrift und Schriftlichkeit. In Hartmut Günther & Otto Ludwig (eds.), *Schrift und Schriftlichkeit/Writing and its use* (= Handbooks of Linguistics and Communication Science, 10.1), 102–121. Berlin, Boston: De Gruyter. DOI: 10.1515/9783110111293.1.2.102

Schmid, Stephan (1997): The Naturalness Differential Hypothesis: cross-linguistic influence and universal preferences in interlanguage phonology and morphology. *Folia Linguistica* 31.3–4: 331–348. DOI: 10.1515/flin.1997.31.3-4.331

Schmidt, Karsten (2014): Morphophonographic regularities in German: the graphematic syllable boundary. *Written Language and Literacy* 17.2: 253–281. DOI: 10.1075/wll.17.2.04sch

Schmidt, Karsten (2016): Der graphematische Satz: Vom Schreibsatz zur allgemeinen Satzvorstellung. *Zeitschrift für germanistische Linguistik* 44.2: 215–256. DOI: 10.1515/zgl-2016-0011

Schmidt, Karsten (2018): *Phonographie und Morphographie im Deutschen. Grundzüge einer wortbasierten Graphematik*. Tübingen: Stauffenburg.

Schmitt, Alfred (1980): *Entstehung und Entwicklung von Schriften*, ed. by Claus Haebler. Köln, Wien: Böhlau.

Schreiber, Gordian (submitted): Visual politeness: remarks on cursivization as found in pre-modern Japanese handbooks on letter writing. In *The idea of writing*, vol. 4. Leiden: Brill.

Schroeder, Klaus-Henning (1981): Schrifttheorie und Konnotation der Schriftzeichen. In Thomas Kotschi (ed.), *Beiträge zur Linguistik des Französischen* (= Tübinger Beiträge zur Linguistik, 154), 123–140. Tübingen: Narr.

Scribner, Sylvia & Michael Cole (1981): *The psychology of literacy*. Cambridge: Harvard University Press.

Sebba, Mark (2007): *Spelling and society: the culture and politics of orthography around the world*. Cambridge: CUP.

Sebba, Mark (2009): Sociolinguistic approaches to writing systems research. *Writing Systems Research* 1.1: 35–49. DOI: 10.1093/wsr/wsp002

Seidenberg, Mark (2011): Reading in different writing systems: One architecture, multiple solutions. In Peggy D. McCardle, Brett Miller, Jun Ren Lee & Ovid J. L. Tzeng (eds.), *Dyslexia across languages: Orthography and the brain–gene–behavior link*, 146–168. Baltimore, MD: Paul H. Brookes.

Seidenberg, Mark & James McClelland (1989): A distributed, developmental model of word recognition and naming. *Psychological Review* 96.4: 523–568. DOI: 10.1037/0033-295X.96.4.523

Seifart, Frank (2006): Orthography development. In Jost Gippert, Nikolaus P. Himmelmann & Ulrike Mosel (eds.), *Essentials of language documentation* (= Trends in Linguistics, 178), 275–299. Berlin, Boston: De Gruyter. DOI: 10.1515/9783110197730.275

Seipel, Wilfried (ed.) (2003): *Der Turmbau zu Babel: Ursprung und Vielfalt von Sprache und Schrift.* Graz: Skira.

Selkirk, Elisabeth O. (1984): *Phonology and syntax: the relation between sound and structure* (= Current Studies in Linguistics, 10). Cambridge: MIT Press.

Seymour, Philip H. K., Mikko Aro & Jane M. Erskine (2003): Foundation literacy acquisition in European orthographies. *British Journal of Psychology* 94.2: 143–174. DOI: 10.1348/000712603321661859

Shany, Michal, Amalia Bar-On & Tami Katzir (2012): Reading different orthographic structures in the shallow pointed Hebrew script: a cross-grade study in elementary school. *Reading and Writing* 25.6: 1217–1238. DOI: 10.1007/s11145-011-9314-y

Share, David L. (2008): On the Anglocentricities of current reading research and practice: the perils of over-reliance on an "outlier" orthography. *Psychological Bulletin* 134.4: 584–616. DOI: 10.1037/00332909.134.4.584

Share, David L. (2014): Alphabetism in reading science. *Frontiers in Psychology* 5: 752. DOI: 10.3389/fpsyg.2014.00752

Share, David L. & Peter T. Daniels (2016): Aksharas, alphasyllabaries, abugidas, alphabets and orthographic depth: Reflections on Rimzhim, Katz and Fowler (2014). *Writing Systems Research* 8.1: 17–31. DOI: 10.1080/17586801.2015.1016395

Sheridan, Heather, Erik D. Reichle & Eyal M. Reingold (2016): Why does removing inter-word spaces produce reading deficits? The role of parafoveal processing. *Psychonomic Bulletin & Review* 23.5: 1543–1552. DOI: 10.3758/s13423-015-0997-y

Shimron, Joseph & David Navon (1981): The distribution of information within letters. *Perception & Psychophysics* 30.5: 483–491. DOI: 10.3758/BF03204845

Shovman, Mark M. & Merav Ahissar (2006): Isolating the impact of visual perception on dyslexics' reading ability. *Vision Research* 56.20: 3514–3525. DOI: 10.1016/j.visres.2006.05.011

Shu, Hua & Richard C. Anderson (1999): Learning to read Chinese: the development of metalinguistic awareness. In Jian Wang, Albrecht W. Inhoff & Hsuan-Chih Chen (eds), *Reading Chinese script: a cognitive analysis*, 1–18. Mahwah: Lawrence Erlbaum.

Shu, Hua, Xi Chen, Richard C. Anderson, Ningning Wu & Yue Xuan (2003): Properties of school Chinese: implications for learning to read. *Child Development* 74.1: 27–47. DOI: 10.1111/1467-8624.00519

Silvia, Paul J. & Christopher M. Barona (2009): Do people prefer curved objects? Angularity, expertise, and aesthetic preference. *Empirical Studies of the Arts* 27.1: 25–42. DOI: 10.2190/EM.27.1.b

Simpson, Ian C., Petroula Mousikou, Juan M. Montoya & Sylvia Defior (2013): A letter visual-similarity matrix for Latin-based alphabets. *Behavior Research Methods* 45.2: 431–439. DOI: 10.3758/s13428-012-0271-4

Sirat, Colette (1988): The material conditions of the lateralization of the ductus. In Derrick de Kerckhove & Charles J. Lumsden (eds.), *The alphabet and the brain: the lateralization of writing*, 173–201. Berlin, Heidelberg: Springer. DOI: 10.1007/978-3-662-01093-8_11

Smalley, William A. (1964): How shall I write this language? In William A. Smalley et al. (eds.), *Orthography studies: articles on new writing systems* (= Helps for Translators, 6), 31–52. London: United Bible Societies.

Smith, Frank (1971): *Understanding reading: a psycholinguistic analysis of reading and learning to read*. Austin: Holt, Rinehart & Winston.

Smith, Janet S. (Shibamoto) (1996): Japanese Writing. In Peter T. Daniels & William Bright (eds.), *The world's writing systems*, 209–217. Oxford: OUP.

Smyth, David (2002): *Thai: an essential grammar*. London, New York: Routledge.

Spillner, Bernd (1974): *Linguistik und Literaturwissenschaft. Stilforschung, Rhetorik, Textlinguistik*. Stuttgart: Wilhelm Kohlhammer.

Spitzmüller, Jürgen (2013): *Graphische Variation als soziale Praxis. Eine soziolinguistische Theorie skripturaler ‚Sichtbarkeit'* (= Linguistik – Impulse & Tendenzen, 56). Berlin, Boston: De Gruyter. DOI: 10.1515/9783110334241

Spitzmüller, Jürgen (2016): Typographie – Sprache als Schriftbild. In Nina-Maria Klug & Hartmut Stöckl (eds.), *Handbuch Sprache im multimodalen Kontext* (= Handbooks of Linguistic Knowledge, 7), 99–120. DOI: 10.1515/9783110296099-005

Spitzmüller, Jürgen (in press): Typographie und soziale Praxis. In Ursula Rautenberg & Anja Voeste (eds.), *Typographie: Disziplinäre Zugänge – Fachliche Konzeptionierungen – Forschungsfragen und Projekte*. Stuttgart: Hiersemann.

Sprenger-Charolles, Liliane, Linda S. Siegel & Philippe Bonnet (1998): Reading and spelling acquisition in French: the role of phonological mediation and orthographic factors. *Journal of Experimental Child Psychology* 68.2: 134–165. DOI: 10.1006/jecp.1997.2422

Sproat, Richard (2000): *A computational theory of writing systems* (= Studies in Natural Language Processing). Cambridge: CUP.

Stalph, Jürgen (1989): *Grundlagen einer Grammatik der sinojapanischen Schrift* (= Veröffentlichungen des Ostasien-Instituts der Ruhr-Universität Bochum, 39). Wiesbaden: Otto Harrassowitz.

Stalph, Jürgen (1996): Das japanische Schriftsystem. In Hartmut Günther & Otto Ludwig (eds.), *Schrift und Schriftlichkeit/Writing and its use* (= Handbooks of Linguistics and Communication Science, 10.2), 1413–1427. Berlin, Boston: De Gruyter. DOI: 10.1515/9783110147445.2.9.1413

Stampe, David (1979): *A dissertation on Natural Phonology*. Bloomington: IULC.

Stampe, David (1985): On the two levels of phonological representations. In Wolfgang U. Dressler & Livia Tonelli (eds.), *Natural Phonology from Eisenstadt: papers on Natural*

*Phonology from the fifth International Phonology Meeting, 25–28 June, 1984*, 131–142. Padova: CLESP.

Stetter, Christian (2011): Ideographie und Alphabetschrift: Zur Ikonizität von Schriften. In Antonio Loprieno, Carsten K. Salis & Birgit Mersmann (eds.), *Bild, Macht, Schrift: Schriftkulturen in bildkritischer Perspektive*, 37–66. Weilerswist: Velbrück Wissenschaft.

Stetter, Christian (2013): Buchstabe und Zeichen. In Ryôsuke Ôhashi & Martin Roussel (eds.), *Buchstaben der Welt – Welt der Buchstaben* (= Morphomata, 12), 87–97. München: Wilhelm Fink.

Stöckl, Hartmut (2004): Typographie: Körper und Gewand des Textes. Linguistische Überlegungen zu typographischer Gestaltung. *Zeitschrift für Angewandte Linguistik* 41: 5–48.

Stokoe, William C. (1960): *Sign language structure. An outline of the visual communication systems of the American Deaf* (= Studies in Linguistics. Occasional Papers, 8). Buffalo: University of Buffalo.

Strätling, Susanne & Georg Witte (2006): Die Sichtbarkeit der Schrift zwischen Evidenz, Phänomenalität und Ikonizität. Zur Einführung in diesen Band. In Susanne Strätling & Georg Witte (eds.), *Die Sichtbarkeit der Schrift* (= Kulturtechnik), 7–20. München: Wilhelm Fink.

Street, Brian V. (1984): *Literacy in theory and practice*. Cambridge: CUP.

Su, I.-Fan, Sin-Ching C. Mak, Lai-Ying M. Cheung & Sam-Po Law (2012): Taking a radical position: evidence for position-specific radical representations in Chinese character recognition using masked priming ERP. *Frontiers in Psychology* 3: 333. DOI: 10.3389/fpsyg.2012.00333

Sung-ik, Ko (2016): An analysis of the Korean alphabet's featural characteristics. *Scripta* 8: 1–24.

Švelch, Jaroslav & Tamah Sherman (2018): "I see your garbage": participatory practices and literary privilege on "Grammar Nazi" Facebook pages in different sociolinguistic contexts. *New Media & Society* 20.7: 2391–2410. DOI: 10.1177/1461444817719087

Sweller, John, Jeroen J. G. van Merrienboer & Fred G. W. C. Paas (1998): Cognitive architecture and instructional design. *Educational Psychology Review* 10.3: 251–296. DOI: 10.1023/A:1022193728205

Szmrecsanyi, Benedikt & Bernd Kortmann (2012): Introduction. In Bernd Kortmann & Benedikt Szmrecsanyi (eds.), *Linguistic complexity: second language acquisition, indigenization, contact*, 6–34. Berlin, Boston: De Gruyter. DOI: 10.1515/9783110229226.6

Taft, Marcus, Xiaoping Zhu & Danling Peng (1999): Positional specificity of radicals in Chinese character recognition. *Journal of Memory and Language* 40.4: 498–519. DOI: 10.1006/jmla.1998.2625

Taha, Haitham Y. (2013): Reading and spelling in Arabic: linguistic and orthographic complexity. *Theory and Practice in Language Studies* 3.5: 721–727. DOI: 10.4304/tpls.3.5.721-727

Taha, Haitham Y. (2016): Deep and shallow in Arabic orthography: new evidence from reading performance of elementary school native Arab readers. *Writing Systems Research* 8.2: 133–142. DOI: 10.1080/17586801.2015.1114910

Taha, Haitham Y., Raphiq Ibrahim & Asaid Khateb (2013): How does Arabic orthographic connectivity modulate brain activity during visual word recognition: an ERP study. *Reading and Writing* 26.2: 292–302. DOI: 10.1007/s10548-012-0241-2

Takagi, Mariko (2014): *Hanzi graphy: a typographic translation between Latin letters and Chinese characters*. Hong Kong: MCCM Creations.

Takagi, Mariko (2019): Kanji/Hiragana/Katakana. In Ben Wittner, Sascha Thoma & Timm

Hartmann (eds.), *Bi-Scriptual: typography and graphic design with multiple script systems,* 281–315. Salenstein: Niggli.

Tamariz, Mónica (2017): Experimental studies on the cultural evolution of language. *Annual Review of Linguistics* 3: 389–407. DOI: 10.1146/annurev-linguistics-011516-033807

Tamariz, Mónica, Seán G. Roberts, J. Isidro Martínez & Julio Santiago (2018): The interactive origin of iconicity. *Cognitive Science* 42.1: 334–349. DOI: 10.1111/cogs.12497

Taylor, Isaac (1883): *The alphabet. An account of the origin and development of letters*, vol. 1*: Semitic alphabet*, vol. 2: *Aryan alphabets*. London: Kegan Paul, Trench.

Terzuolo, Carlo A. & Paolo Viviani (1980): Determinants and characteristics of motor patterns used for typing. *Neuroscience* 5: 1085–1103. DOI: 10.1016/0306-4522(80)90188-8

Thomassen, Arnold J.W.M. & Hein J.C.M. Tibosch (1991): A quantitative model of graphic production. In Jean Requin & George E. Stelmach (eds.), *Tutorials in motor neuroscience*, 269–281. Dordrecht: Springer. DOI: 10.1007/978-94-011-3626-6_22

Thomassen, Arnold J. W. M., Ruud G. J. Meulenbroek & Marion P. E. Hoofs (1992): Economy and anticipation in graphic stroke sequences. *Human Movement Science* 11.1–2: 71–82. DOI: 10.1016/0167-9457(92)90051-C

Tolchinsky-Landsmann, Liliana & Iris Levin (1985): Writing in preschoolers: an age-related analysis. *Applied Psycholinguistics* 6.3: 319–339. DOI: 10.1017/S0142716400006238

Tolchinsky, Liliana (2003): *The cradle of culture and what children know about writing and numbers before being taught* (= The Developing Mind Series). Mahwah, NJ: Lawrence Erlbaum.

Tong, Xiuhong, Xiuli Tong & Catherine McBride (2017): Radical sensitivity is the key to understanding Chinese character acquisition in children. *Reading and Writing* 30.6: 1251–1265. DOI: 10.1007/s11145-017-9722-8

Totereau, Corinne, Marie-Geneviève Thevenin & Michel Fayol (1997): The development of the understanding of number morphology in written French. In Charles E. Perfetti, Laurence Rieben & Michel Fayol (eds.), *Learning to spell: research, theory, and practice across languages*, 97–114. Mahwah: Lawrence Erlbaum.

Treiman, Rebecca, Brett Kessler & Tatiana C. Pollo (2006): Learning about the letter name subset of the vocabulary: evidence from US and Brazilian preschoolers. *Applied Psycholinguistics* 27.2: 211–227. DOI: 10.1017/S0142716406060255

Treiman, Rebecca, Jeremy Cohen, Kevin Mulqueeny, Brett Kessler & Suzanne Schechtman (2007): Young children's knowledge about printed names. *Child Development* 78.5: 1458–1471. DOI: 10.1111/j.1467-8624.2007.01077.x

Treiman, Rebecca & Brett Kessler (2014): *How children learn to write words*. Oxford: OUP.

Treiman, Rebecca, Kevin Mulqueeny & Brett Kessler (2015): Young children's knowledge about the spatial layout of writing. *Writing Systems Research* 7.2: 235–244. DOI: 10.1080/17586801.2014.924386

Trudgill, Peter (2000): *Sociolinguistics: an introduction to language and society*, 4th edition. London: Penguin Books.

Tschichold, Jan (1981): *Formenverwandlungen der &-Zeichen*. Dresden: VEB Verlag der Kunst.

Tversky, Barbara (1995): Cognitive origins of graphic productions. In Francis T. Marchese (ed.), *Understanding images: finding meaning in digital imagery*, 29–53. Berlin: Springer.

Underwood, N. Roderic & George W. McConkie (1985): Perceptual span for letter distinctions during reading. *Reading Research Quarterly* 20.2: 153–162. DOI: 10.2307/747752

Unseth, Peter (2005): Sociolinguistic parallels between choosing scripts and languages. *Written Language and Literacy* 8.1: 19–42. DOI: 10.1075/wll.8.1.02uns

Vachek, Josef (1989): Some remarks on writing and phonetic transcription. In Josef Vachek, Philip A. Luelsdorff (eds.), *Written language revisited*, 1–8. Amsterdam, Philadelphia: John Benjamins.

van Drempt, Nadege, Annie McCluskey & Natasha A. Lannin (2011): A review of factors that influence adult handwriting performance. *Australian Occupational Therapy Journal* 58.5: 321–328. DOI: 10.1111/j.1440-1630.2011.00960.x

van Gemmert, Arend W. A. & Jose L. Contreras-Vidal (2015): Graphonomics and its contribution to the field of motor behavior: a position statement. *Human Movement Science* 43: 165–168. DOI: 10.1016/j.humov.2015.08.017

Velan, Hadas & Ram Frost (2007): Cambridge University versus Hebrew University: the impact of letter transposition on reading English and Hebrew. *Psychonomic Bulletin & Review* 14.5: 913–918. DOI: 10.3758/BF03194121

Velasco, Carlos, Andy T. Woods, Sarah Hyndman & Charles Spence (2015): The taste of typeface. *i-Perception* 6.4: 1–10. DOI: 10.1177/2041669515593040

Venezky, Richard L. (1977): Principles for the design of practical writing systems. In Joshua Fishman (ed.), *Advances in the creation and revision of writing systems* (= Contributions to the Sociology of Language, 8), 37–54. Berlin, Boston: De Gruyter. DOI: 10.1515/9783110807097-004

Venezky, Richard L. (2004): In search of the perfect orthography. *Written Language and Literacy* 7.2: 139–163. DOI: 10.1075/wll.7.2.02ven

Vennemann, Theo (1988): *Preference laws for syllable structure and the explanation of sound change with special reference to German, Germanic, Italian, and Latin*. Berlin, Boston: De Gruyter. DOI: 10.1515/9783110849608

Voeste, Anja (2004): Graphotaktik. In Jörg Riecke (ed.), *Einführung in die historische Textanalyse*, 183–202. Göttingen: Vandenhoeck & Ruprecht.

Voß, Viola (2003): *Schrifttypologie und das japanische Schriftsystem* (= Berliner Beiträge zur Linguistik, 2). Berlin: Weißensee Verlag.

Waller, Robert H. W. (1991): Typography and discourse. In Rebecca Barr, Michael L. Kamil, Paul D. Pearson & Peter Rosenthal (eds.), *Handbook of reading research*, vol. 2, 341–380. New York: Longman.

Wang, Fang-yü (1958): *Introduction to Chinese cursive script*. New Haven: Institute of Far Eastern Languages, Yale University.

Wathabunditkul, Suphawut (2003): *Spacing in the Thai language*. <http://www.thai-language.com/ref/spacing> (September 22nd, 2020).

Watkins, Justin (2009): Burmese. In Keith Brown & Sarah Ogilvie (eds.), *Concise encyclopedia of languages of the world*, 170–175. Amsterdam: Elsevier.

Watt, W. C. (1975): What is the proper characterization of the alphabet? I: Desiderata. *Visible Language* 9: 293–327.

Watt, W. C. (1979): Iconic equilibrium. *Semiotica* 28.1–2: 31–62. DOI: 10.1515/semi.1979.28.1-2.31

Watt, W. C. (1980): What is the proper characterization of the alphabet? – II: Composition. *Ars Semeiotica* 3.1: 3–46.

Watt, W. C. (1981): What is the proper characterization of the alphabet? – III: Appearance. *Ars Semeiotica* 4.3: 269–313.

Watt, W. C. (1983a): Grade der Systemhaftigkeit: Zur Homogenität der Alphabetschrift. *Zeitschrift für Semiotik* 5.4: 371–399.

Watt, W. C. (1983b): Mode, modality, and iconic evolution. In Tasso Borbé (ed.), *Semiotics Unfolding. Proceedings of the Second Congress of the International Association for Semiotic Studies, Vienna, July 1979*, vol. III, 1543–1550. Berlin, Boston: De Gruyter. DOI: 10.1515/9783110869897

Watt, W. C. (1988): What is the proper characterization of the alphabet? IV: Union. *Semiotica* 70.3–4: 199–241. DOI: 10.1515/semi.1988.70.3-4.199

Watt, W. C. (1989): Getting writing right. *Semiotica* 75.3–4: 279–315. DOI: 10.1515/semi.1989.75.3-4.279

Watt, W. C. (1994a): Writing systems: introduction. In W. C. Watt (ed.), *Writing systems and cognition* (= Neuropsychology and Cognition, 6), 89–114. Dordrecht: Springer. DOI: 10.1007/978-94-015-8285-8_6

Watt, W. C. (1994b): Curves as angles. In W. C. Watt (ed.), *Writing systems and cognition* (= Neuropsychology and Cognition, 6), 215–246. Dordrecht: Springer. DOI: 10.1007/978-94-015-8285-8_11

Watt, W. C. (1998): The old-fashioned way. *Semiotica* 122.1–2: 99–138. DOI: 10.1515/semi.1998.122.1-2.1

Watt, W. C. (2002): Mere writing. *Semiotica* 141: 453–476. DOI: 10.1515/semi.2002.064

Watt, W. C. (2015): What is the proper characterization of the alphabet? VII: Sleight of hand. *Semiotica* 207: 65–88. DOI: 10.1515/sem-2015-0064

Waugh, Linda R. (1982): Marked and unmarked: a choice between unequals in semiotic structure. *Semiotica* 38.3–4: 299–318. DOI: 10.1515/semi.1982.38.3-4.299

Weder, Mirjam (2016): *Orthographische Varianten in der literalen Praxis: empirische Untersuchung des Usus, der individuellen Repräsentation und der Wirkung auf den Schreibprozess* (= Basler Studien zur deutschen Sprache und Literatur, 95). Tübingen: Narr Francke Attempto.

Wehde, Susanne (2000): *Typographische Kultur. Eine zeichentheoretische und kulturgeschichtliche Studie zur Typographie und ihrer Entwicklung* (= Studien und Texte zur Sozialgeschichte der Literatur, 69). Berlin, Boston: De Gruyter. DOI: 10.1515/9783110945799

Weingarten, Rüdiger (2011): Comparative graphematics. *Written Language and Literacy* 14.1: 12–38. DOI: 10.1075/wll.14.1.02wei

Weth, Constanze & Kasper Juffermans (2018): Introduction: the tyranny of writing in language and society. In Constanze Weth & Kasper Juffermans (eds.), *The tyranny of writing: ideologies of the written word*, 1–19. London, New York: Bloomsbury.

Whaley, Lindsay J. (1997): *Introduction to typology: The unity and diversity of language*. Thousand Oaks, London: SAGE.

White, Sarah J., Rebecca L. Johnson, Simon P. Liversedge & Keith Rayner (2008): Eye movements when reading transposed text: the importance of word-beginning letters. *Journal of Experimental Psychology: Human Perception and Performance* 34.5: 1261–1276. DOI: 10.1037/0096-1523.34.5.1261

Whitney, Carol (2001): How the brain encodes the order of letters in a printed word: the SERIOL model and selective literature review. *Psychonomic Bulletin & Review* 8.2: 221–243. DOI: 10.3758/BF03196158

Wiebelt, Alexandra (2003): Die Entwicklung der Symmetrie in der Schrift: Wie Objektkonstanz die Genese von Buchstabenformen beeinflusst. *Linguistische Berichte* 195: 295–323.

Wiebelt, Alexandra (2004a): *Symmetrie bei Schriftsystemen: ein Lesbarkeitsproblem* (= Linguistische Arbeiten, 488). Berlin, Boston: De Gruyter. DOI: 10.1515/9783110919721

Wiebelt, Alexandra (2004b): Do symmetrical letter pairs affect readability? A cross-linguistic examination of writing systems with special reference to runes. *Written Language and Literacy* 7.2: 275–304. DOI: 10.1075/wll.7.2.07wie

Wiese, Richard (1996): *The phonology of German* (= The Phonology of the World's Languages). Oxford: OUP.

Will, Udo, Guido Nottbusch & Rüdiger Weingarten (2006): Linguistic units in word typing.

Effects of word presentation modes and typing delay. *Written Language and Literacy* 9.1: 153–176. DOI: 10.1075/wll.9.1.10wil

Willberg, Hans P. & Friedrich Forssman (2010): *Lesetypografie*, 4th edition. Mainz: Hermann Schmidt.

Willows, Dale M. & Megan Terepocki (1993): The relation of reversal errors to reading disabilities. In Dale M. Willows, Richard S. Kruk & Evelyne Corcos (eds.), *Visual processes in reading and reading disabilities*, 31–56. Hillsdale: Lawrence Erlbaum Associates.

Willows, Dale M. & Esther Geva (1995): What is visual in orthographic processing? In: Virginia W. Berninger (ed.), *The varieties of orthographic knowledge* (= Neuropsychology and Cognition, 11), 355–376. Springer: Dordrecht. DOI: 10.1007/978-94-011-0385-5_11

Wimmer, Heinz & Heinz Mayringer (2002): Dysfluent reading in the absence of spelling difficulties: a specific disability in regular orthographies. *Journal of Educational Psychology* 94: 272–277. DOI: 10.1037/0022-0663.94.2.272

Wing, Alan M. (1979): Variability in handwritten characters. *Visible Language* 13.3: 283–298.

Winskel, Heather (2009): Reading in Thai: the case of misaligned vowels. *Reading and Writing* 22.1: 1–24. DOI: 10.1007/s11145-007-9100-z

Winskel, Heather (2010): Spelling development in Thai children. *Journal of Cognitive Science* 11.1: 7–35. DOI: 10.17791/jcs.2010.11.1.7

Winskel, Heather (2014): Learning to read and write in Thai. In Heather Winskel & Prakash Padakannaya (eds.), *South and Southeast Asian psycholinguistics*, 171–178. Cambridge: CUP. DOI: 10.1017/CBO9781139084642.019

Winskel, Heather (2016): Insights into reading processes through investigating diversity. *Australian Journal of Psychology* 69.3: 151–161. DOI: 10.1111/ajpy.12136

Winskel, Heather, Ralph Radach & Sudaporn Luksaneeyanawin (2009): Eye movements when reading spaced and unspaced Thai and English: a comparison of Thai-English bilinguals and English monolinguals. *Journal of Memory and Language* 61.3: 339–351. DOI: 10.1016/j.jml.2009.07.002

Winskel, Heather & Kanyarat Iemwanthong (2010): Reading and spelling acquisition in Thai children. *Reading and Writing* 23.9: 1021–1053. DOI: 10.1007/s11145-009-9194-6

Winskel, Heather, Manuel Perea & Theeraporn Ratitamkul (2012): On the flexibility of letter position coding during lexical processing: evidence from eye movements when reading Thai. *The Quarterly Journal of Experimental Psychology* 65.8: 1522–1536. DOI: 10.1080/17470218.2012.658409

Winskel, Heather & Manuel Perea (2014): Can parafoveal-on-foveal effects be obtained when reading an unspaced alphasyllabic script (Thai)? *Writing Systems Research* 6.1: 94–104. DOI: 10.1080/17586801.2013.843440

Wolfe, Jeremy M. (2000): Visual attention. In Karen K. De Valois (ed.), *Seeing*, 335–386. London: Academic Press. DOI: 10.1016/B978-012443760-9/50010-6

Wollscheid, Sabine, Jørgen Sjaastag & Catherine Tømte (2016): The impact of digital devices vs. pen(cil) and paper on primary school students' writing skills – a research review. *Computers & Education* 95: 19–35. DOI: 10.1016/j.compedu.2015.12.001

Wong, Kimberly, Frempongma Wadee, Gali Ellenblum & Michael McCloskey (2018): The devil's in the g-tails: deficient letter-shape knowledge and awareness despite massive visual experience. *Journal of Experimental Psychology: Human Perception and Performance* 44.9: 1324–1335. DOI: 10.1037/xhp0000532

Wood, Margaret M. (1929): Latinizing the Turkish alphabet: a study in the introduction of a cultural change. *American Journal of Sociology* 35.2: 194–203.

Wurzel, Wolfgang U. (1988): The structuralist heritage in natural morphology. *Linguistische Studien* A 188: 99–116.

Wurzel, Wolfgang U. (1994): Natural Morphology. In R. E. Asher & J. M. Y. Simpson (eds.), *The encyclopedia of language and linguistics*, 2590–2598. Oxford: Pergamon Press.

Xiao, Wen & Rebecca Treiman (2012): Iconicity of simple Chinese characters. *Behavior Research Methods* 44.4: 954–960. DOI: 10.3758/s13428-012-0191-3

Xu, Min (2015): Character amnesia. In Rint Sybesma (ed.), *Encyclopedia of Chinese language and linguistics*. Leiden: Brill. DOI: 10.1163/2210-7363_ecll_COM_000225

Yakup, Mahire, Wayit Abliz, Joan Sereno & Manuel Perea (2014): How is letter position coding attained in scripts with position-dependent allography? *Psychonomic Bulletin & Review 21*.6: 1600–1606. DOI: 10.3758/s13423-014-0621-6

Yan, Guoli, Hongjie Tian, Xuejun Bai & Keith Rayner (2006): The effect of word and character frequency on the eye movements of Chinese readers. *British Journal of Psychology* 97.2: 259–268. DOI: 10.1348/000712605X70066

Yan, Ming, Reinhold Kliegl, Eike M. Richter, Antje Nuthmann & Hua Shu (2010): Flexible saccade-target selection in Chinese reading. *Quarterly Journal of Experimental Psychology* 63.4: 705–725. DOI: 10.1080/17470210903114858

Yan, Zhenjiang (2002): Der geheime Phono- und Eurozentrismus des Redens von Schrift. In Erika Greber, Konrad Ehlich & Jan-Dirk Müller (eds.), *Materialität und Medialität von Schrift*, 151–164. Bielefeld: Aisthesis.

Yao, Yun (2011): Interword spacing effects on reading Mandarin Chinese as a second language. *Writing Systems Research* 3.1: 23–40. DOI: 10.1093/wsr/wsr009

Yen, Miao-Hsuan, Ralph Radach, Ovid J.-L. Tzeng, Daisy L. Hung & Jie-Li Tsai (2009): Early parafoveal processing in reading Chinese sentences. *Acta Psychologica* 131.1: 24–33. DOI: 10.1016/j.actpsy.2009.02.005

Yen, Miao-Hsuan, Ralph Radach, Ovid J.-L. Tzeng & Jie-Li Tsai (2012): Usage of statistical cues for word boundary in reading Chinese. *Reading and Writing* 25.5: 1007–1029. DOI: 10.1007/s11145-011-9321-z

Yu, Xiaoli (2001): Zur Komplexität chinesischer Schriftzeichen. *Göttinger Beiträge zur Sprachwissenschaft* 5: 121–129.

Yunick, Stanley (2000): Linguistics, TESL, and language planning in Micronesia. *Studies in the Linguistic Sciences* 30.1: 183–200.

Zang, Chuanli, Feifei Liang, Xuejun Bai, Guoli Yan & Simon P. Liversedge (2013): Interword spacing and landing position effects during Chinese reading in children and adults. *Journal of Experimental Psychology: Human Perception & Performance* 39.3: 720–734. DOI: 10.1037/a0030097

Zemlock, Deborah, Sophia Vinci-Booher & Karin H. James (2018): Visual-motor symbol production facilitates letter recognition in young children. *Reading and Writing* 31.6: 1255–1271. DOI: 10.1007/s11145-018-9831-z

Zhang, Haiwai (2014): A review of stroke order in hanzi handwriting. *Journal of the European Confederation of Language Centres in Higher Education* (*CercleS*) 4.2: 423–440. DOI: 10.1515/cercles-2014-0022

Zhang, Lan, Li Yin & Rebecca Treiman (2016): Chinese children's early knowledge about writing. *British Journal of Developmental Psychology* 35.3: 349–358. DOI: 10.1111/bjdp.12171

Zhou, Wei, Aiping Wang, Hua Shu, Reinhold Kliegl & Ming Yan (2018): Word segmentation by alternating colors facilitates eye guidance in Chinese reading. *Memory & Cognition* 46.5: 729–740. DOI: 10.3758/s13421-018-0797-5

Ziefle, Martina (2002): *Lesen am Bildschirm. Eine Analyse visueller Faktoren* (= Internationale Hochschulschriften, 375). Münster, New York: Waxmann.

Ziegler, Johannes C. & Usha Goswami (2005): Reading acquisition, developmental dyslexia, and skilled reading across languages: a psycholinguistic grain size theory. *Psychological Bulletin* 131.1: 3–29. DOI: 10.1037/0033-2909.131.1.3

Ziegler, Johannes C., Daisy Bertrand, Dénes Tóth, Valéria Csépe, Alexandra Reis, Luís Faísca, Nina Saine, Heikki Lyytinen, Anniek Vaessen & Leo Blomert (2010): Orthographic depth and its impact on universal predictors of reading: a cross-language investigation. *Psychological Science* 21.4: 551–559. DOI: 10.1177/0956797610363406

Zifonun, Gisela, Ludger Hoffmann, Bruno Strecker, Joachim Ballweg, Ursula Braußе, Eva Breindl, Ulrich Engel, Helmut Frosch, Ursula Hoberg & Klaus Vorderwülbecke (1997): *Grammatik der deutschen Sprache.* 3 Bde. (= Schriften des Instituts für Deutsche Sprache, 7). Berlin, Boston: De Gruyter. DOI: 10.1515/9783110872163

Zinn, Ernst (1950/51): Schlangenschrift. *Archäologischer Anzeiger* 65/66: 1–36.

Zwicky, Arnold (1973): The strategy of generative phonology. *Ohio State University Working Papers in Linguistics* 14: 85–99.

# VIII Index

## A

# D

# E



# F

## G

## H

# M

# N

# O



# P

# Q



# R

# T



# U

## V



## W



## X



## Y



## Z