DIMITRIOS MELETIS

University of Zurich

OPEN QUESTIONS IN THE CROSS-LINGUISTIC CONCEPTION OF THE GRAPHEME

↗ Linguistic value, (supra)segmentality, and cognitive reality

SCRIPTA 2023

National University of Mongolia, July 10, 2023

STARTING POINT

* proposal of defining the grapheme as a **universal basic unit of writing** (MELETIS 2019)

– every writing system has a basic unit (whether defined functionally or formally)

– as there are cross-linguistic concepts in other linguistic domains (phoneme, morpheme), there should be ones in the study of writing, too, to facilitate comparisons between writing systems (CF. MELETIS 2020)

– Or should there not? Is writing – as a secondary semiotic system – different somehow?

– Given the diversity of writing systems, is it feasible to conceptualize the grapheme as a looser comparative concept? (CF. HASPELMATH 2010)

– the structuralist tools used to describe and capture other domains of language can also be applied to writing (*CONTRA* DANIELS 1991)

THREE CRITERIA

1. **distinctiveness criterion**
   - needs to be semantically, specifically lexically distinctive
   - important for the identification of allographs vs. graphemes

2. **linguistic value criterion**
   - needs to relate to a linguistic value (unit, information, structure) somehow
   - relation does not need to be stable or present in every context, it just needs to exist within the system and its use
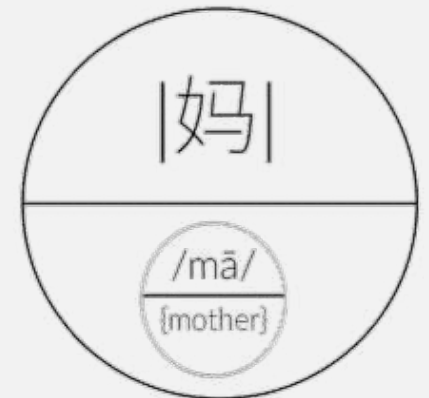
3. **minimality criterion**
   - graphemes are the smallest units for which criteria (1) and (2) apply
   - thus, for example, syllable blocks in Hangul are not graphemes but grapheme combinations; the graphemes are graphetically subsegmental

phonographic grapheme
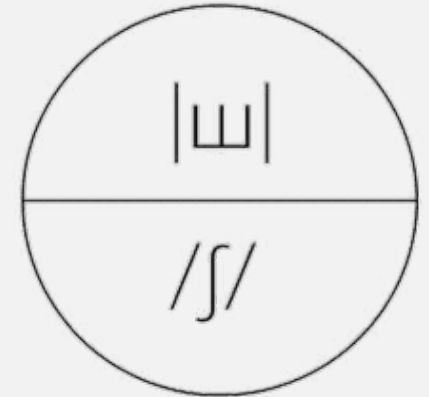
morphographic grapheme

THREE CRITERIA

	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| (has-ta) 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/	+

(MELETIS 2020: 104)

PROBLEMS AND CHALLENGES

1. What about polyvalence?

2. What about elements that can only be analyzed in context? Elements that do not have a phonological representation? Does a purely segmental analysis have its limits?

3. What about the cognitive reality of graphemes and graphematic elements? Do description and cognitive reality align?

   – And is that necessary?

CRITICISM

The less unit-ness of grapheme in the Japanese writing system
(OKADA 2021)

"[…] one should aware [sic] that no graph(eme)s work for a single functionality; but represent several functionality [sic] partially and at the same time" (Slide 2)
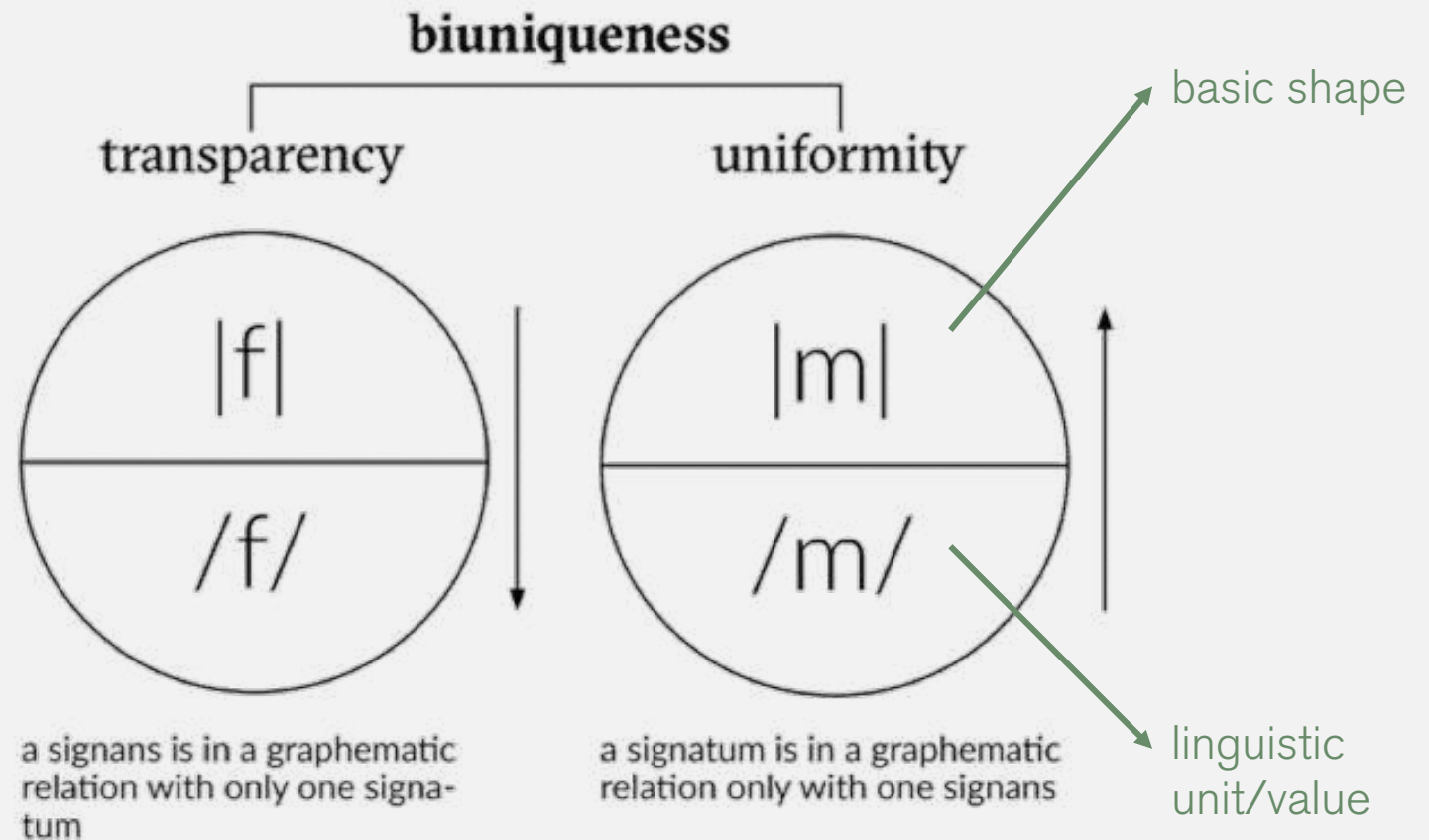
"[…] we find the best graphematic theory is one that accounts for any possibilities but rejects any impossibilities" (Slide 25)

STARTING POINT / LINGUISTIC VALUE / SUPRASEGMENTALITY / COGNITIVE REALITY / OUTLOOK

LACK OF BIUNIQUENESS

semiotic criteria for the description of writing systems
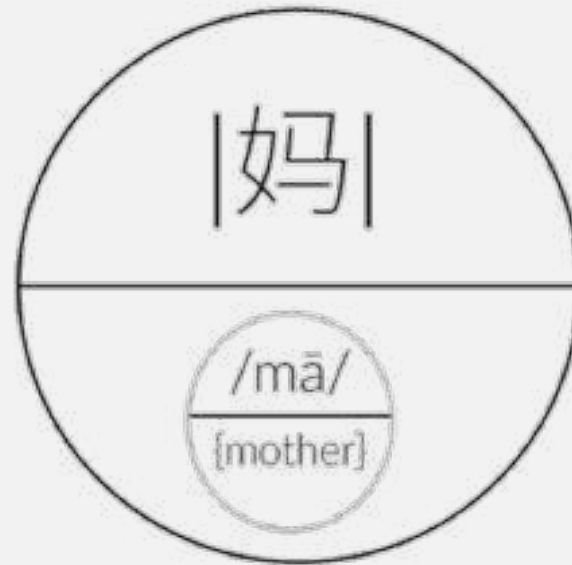
(MELETIS 2020)

in many writing systems, there is a certain **lack of transparency** and/or uniformity

biuniqueness

transparency   uniformity

basic shape

|f| |m|
―――   ―――
/f/   /m/

a signans is in a graphematic relation with only one signatum

a signatum is in a graphematic relation only with one signans

linguistic unit/value

COMPOSITIONAL TRANSPARENCY

compositional
transparency

|妈|

/mā/

{mother}

the sum of the graphematic
value of the grapheme equals
the graphematic value of the
whole grapheme

OKURIGANA: NO PROBLEM

torishimaru
(1) 取りしまる
(2) 取り締まる
(3) 取り締る
(4) 取締る

(OKADA 2021: SLIDE 15, CF. ALSO HONDA 2009)

LINGUISTIC VALUE REVISITED

"In order to identify a unit as a grapheme, it is not necessary for it to refer to only one linguistic unit, and its linguistic reference does not need to be stable. It is only imperative that it has a linguistic value **in all contexts** in which it is used." (CF. MELETIS 2019: 36, EMPHASIS ADDED)

– the direction of analysis is *grapheme → linguistic unit* (or *writing → language*)

– polyvalence, per se, is no problem, as is the context-dependence of specific values; each grapheme still has a default value (CF. NEEF 2005)

– however, against the background that it is necessary in all contexts, what is 'linguistic value', really?

EXTENSION / DETERMINATION

"[. . .] the mechanisms of extension are both simple and universal, and must therefore be intuitively obvious to human beings once the crucial breakthrough of **associating signs with words** has been achieved. The two most basic mechanisms are phonetic extension and semantic extension. These mechanisms are powerful, providing the flexibility and combinatorial power needed to represent spoken language. They carry with them, however, an inevitable disadvantage: they lead to polyvalency in graphic representation, increasing the possibility for ambiguity and confusion, and thus increasing the cognitive load on the process of reading. Techniques of disambiguation naturally follow, including semantic determination, phonetic determination, and graphic modification."

(HANDEL 2019: 309–310, MY EMPHASIS)

DETERMINATIVES

"Semantic determinatives are placed adjacent to a logographic, logophonetic, or phonetic spelling of a word; they disambiguate between alternative orthographic values—not necessarily only between alternative words spelled out by the same sign or sequence of signs, but more specifically between alternative orthographic values, of whatever type, of the same sign or sequence of signs."

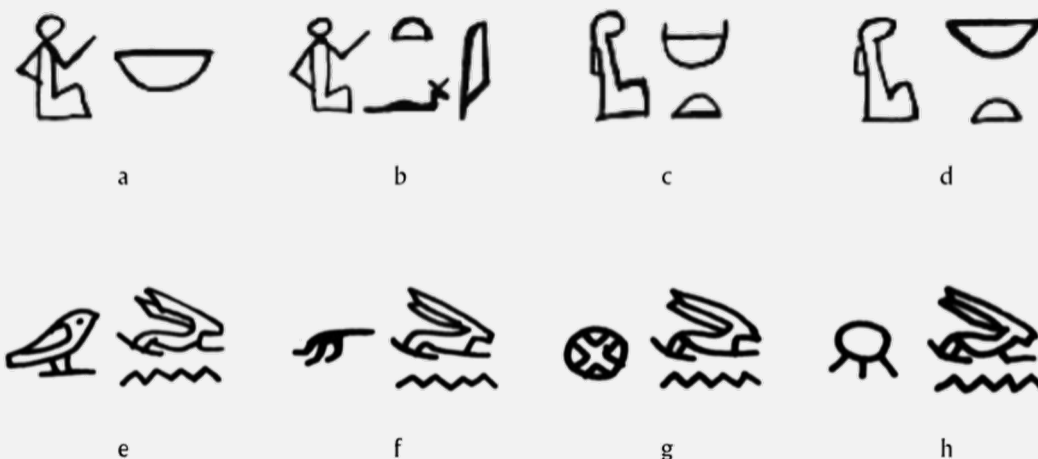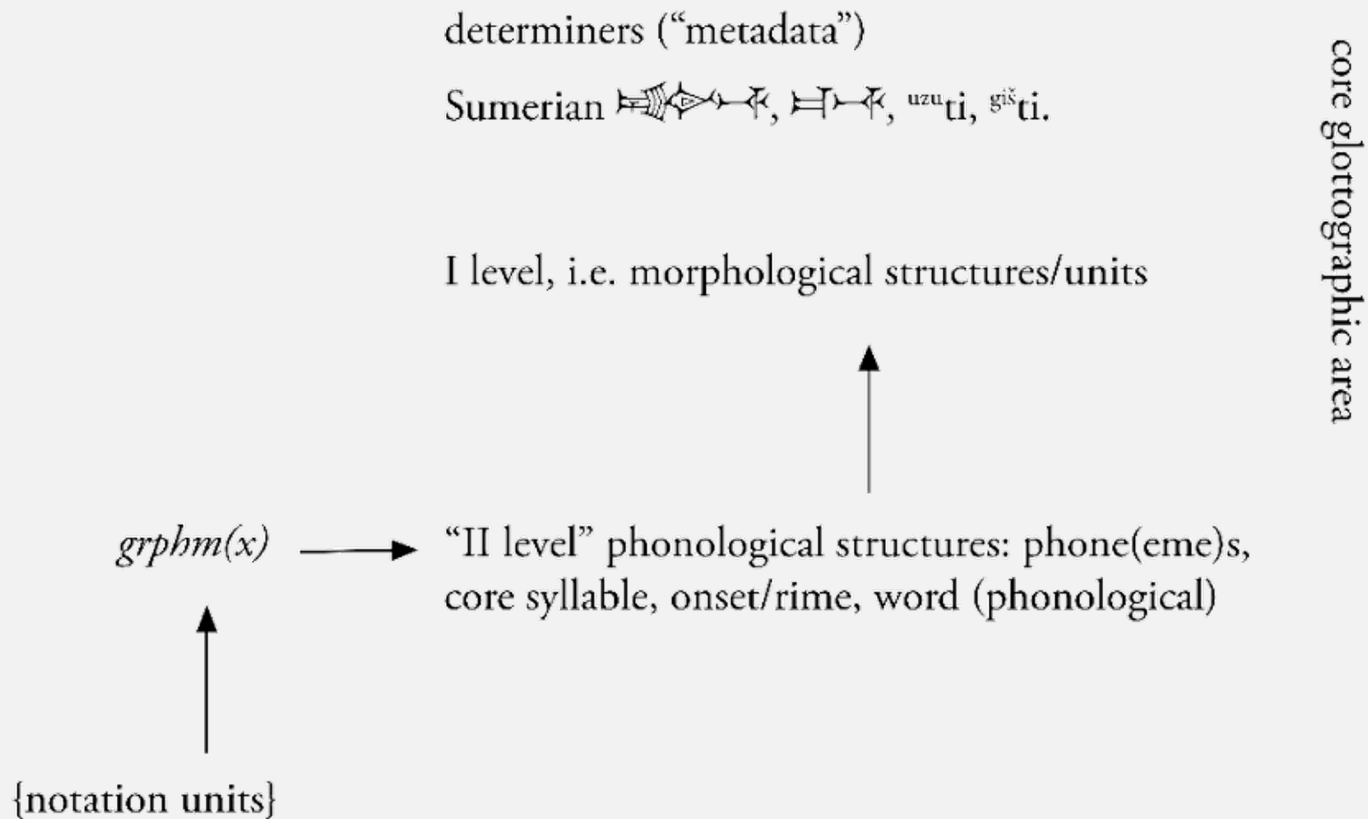(MORA–MARÍN 2008: 195–196)

DETERMINATIVES

Figure 1. Cross-script comparative examples. (a) Egyptian spelling: jt[MAN] for *jt* 'father'; (b) Egyptian spelling: jt[MAN] for *jt* 'father'; (c) Egyptian spelling: 'm-t[WOMAN] for '*m.t* 'wife'; (d) Egyptian spelling: nb-t[WOMAN] for *nb.t* 'lady, mistress'; (e) Egyptian spelling: w-n[EVIL.BIRD] 'mistake'; (f) Egyptian spelling: w-n[LOCK.OF.HAIR] 'bald'; (g) Egyptian spelling: w-n[CITY] 'Hermopolis'; (h) Egyptian spelling: w-n[SUN.RAYS] 'light'.

DETERMINATIVES

determinatives

as metadata?

determiners ("metadata")

Sumerian [cuneiform signs], [cuneiform signs], ^{uzu}ti, ^{giš}ti.

I level, i.e. morphological structures/units

grphm(x) → "II level" phonological structures: phone(eme)s, core syllable, onset/rime, word (phonological)

{notation units}

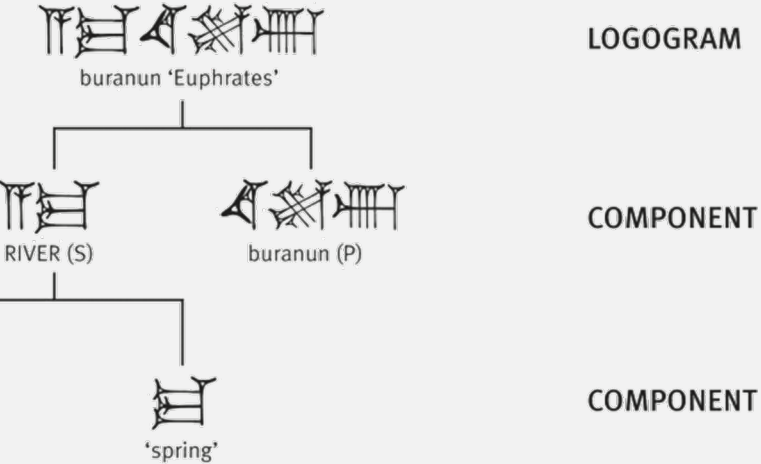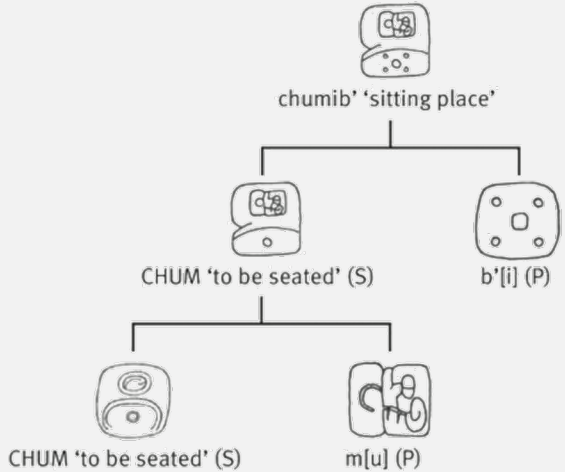core glottographic area

(RIZZA 2021: 33)

DETERMINATIVES: LAYERING
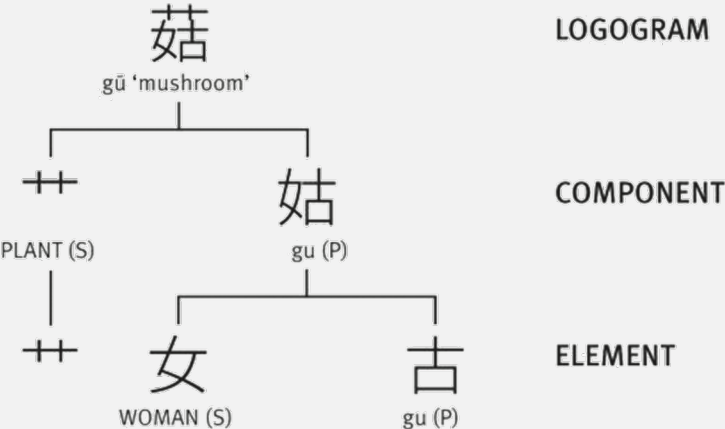
(ADAPTED FROM TRANTER 2013)

DETERMINATIVES

– no problem when they are a part of graphemes; consider the **phonological components** (phonetics) and **semantic components** (radicals) in Chinese graphemes

– How are they conceptualized, however, when they are not graphetically incorporated into graphemes but are their own graphetic segments?

DETERMINATIVES

Are determinatives just **repurposed graphemes**, e.g., graphemes used in a different context where they – from a suprasegmental point of view – become null/zero?

– similar to silent/mute letters in alphabetic writing systems? (CF. ROYER, SPINELLI & FERRAND 2005)

– maybe, but this then only takes into account their graphematic status, not their specific graphematic **function** as determinatives

– but: Are there determinatives that do not derive from (default) graphemes?

– solution (?): determinatives retain part of the function of the graphemes they derive from and are (dependent) graphematic elements, like the subsegmental components in Chinese – in other words, they retain partial functionalities (CF. OKADA 2021)

NULL AFFIXATION?

"a process by which new words are formed by adding an affix which happens to be phonologically null"

"Zero morphology, null morphology, or ø-exponence refers to a particular type of hypothesized object which according to some accounts natural languages have: morphological units—typically, **bound** morphemes—which make a grammatical or **semantic** contribution without directly introducing any phonological information." (DAHL & FÁBREGAS 2018, MY EMPHASIS)

Thus, (graphetically) **segmental bound graphematic elements**?
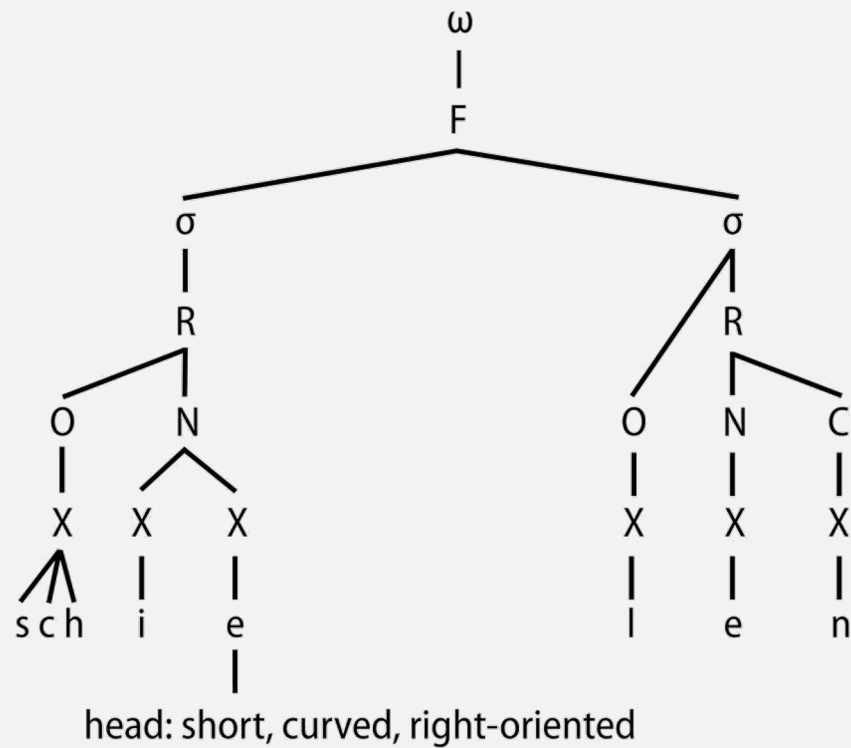
GRAPHEMATIC FUNCTION

– **other segmental bound graphematic elements:** |c| in German is not a grapheme, but part of the complex grapheme <ch> /x/, the grapheme combination <sch> /ʃ/, and syllabically conditioned combinations such as <ck> /k/

– such elements as **value-contributing graphematic elements** vs. determinatives as **value-distinguishing graphematic elements**

EPILOGUE: THE PITFALL OF SEGMENTALITY

Is the core unit of writing the written **word**, and is the core intermodal correspondence the one between written and spoken **words**? This would make segmental grapheme-phoneme-correspondences **epiphenomenal** (CF. STETTER 2011; SCHMIDT 2018). It also reduces the relevance of the questions asked here as determinatives would just be one constitutive part of a unit that is analyzed holistically.

From this perspective, every part of a written word contributes to the value of the entire word.

20

A SUPRASEGMENTAL APPROACH (FOR ALPHABETS)

graphematic word

graphematic foot

graphematic syllables

constituents of the syllable:
O = onset
N = nucleus; R = rime
C = coda

skeletal positions (*graphemes*)

letters

letter features for |e|

reminiscent of the layering approach?

(MELETIS 2019: 31, ADAPTED FROM BERG, PRIMUS & WAGNER 2016: 351)

COGNITIVE REALITY AND _–EME_-STATUS

"The next and necessary step is to evaluate this conception of the grapheme in the light of available data from the processing of writing. Are graphemes as defined here psychologically real? What role do they play in the acquisition of reading and writing processes and are they relevant once these processes are mastered?" (MELETIS 2019: 43)

– definition has already been used in cognitive works (LACHMANN & BERGSTRÖM 2023)

UNCONSCIOUS VS. CONSCIOUS?

"Emic terminology relates to a property of the human mind. It applies to language, and to culture and many aspects of human behavior, because human brains have evolved over the myradennia in a way that produces such behavior. Even if I knew how or why, this would not be the place to go into it; it seems to me simply self-evident. […] Therefore the emic system should not be expected to apply to writing, because writing is a conscious invention. […] I submit it is foolish to try to apply a model developed to account for unconsciously evolved phenomena of human behavior to phenomena that are consciously devised." (DANIELS 1991: 534)

WHY IS THIS INACCURATE?

– literacy affords cognitive tools of conceptualizing and (consciously) thinking about language ↗ (meta)linguistic awareness

– through psychologically salient and real units such as the grapheme and the graphematic word, it has helped us shape cognitive 'fictions' that, due to their suitability and descriptive success, have become cognitive reality – this includes concepts such as the **phoneme** and the **word** (CF. DAVIDSON 2019)

– that way, writing can be seen as a prerequisite or at least driver of linguistics, which manifests itself in **scripticism** or the **written language bias** (CF. LINELL 2005)

– literacy may have been conscious for the first inventors but through the thousands of years of subsequent use writing has become very much a (natural) system **resembling other linguistic systems** (MELETIS 2020) and users do not need to be conscious of its systematics when using it

WHY THE GRAPHEME?

– What is its (not linguistic, but) epistemological value?

  – the results of trying to formulate a cross-linguistically applicable grapheme definition are not what is most important; the **process** of getting there (or trying to) is

  – this already involves the real and interesting **conceptual and comparative work** that requires finding the commonalities between writing systems

  – the question is at which level they are located, and what, in turn, the uncovered universality and diversity in the structural makeup of writing systems may reveal about how we **process** them

OUTLOOK

– we need to further investigate the graphematic status of

  – **punctuation**

  – the semasiographic use of graph(et)ic material not classified as writing, such as, nowadays, **emojis** (CF. DÜRSCHEID & MELETIS 2019)

  – the idiosyncratic functions of frequent graphematic combinations including **digraphs** (CF. OSTERKAMP & SCHREIBER 2021), **complex graphemes** (CF. REINKEN 2022), as well as irregular **ligatures** (e.g., in abugidic systems)

REFERENCES

Berg, Kristian, Beatrice Primus & Lutz Wagner. 2016. Buchstabenmerkmal, Buchstabe, Graphem. In Ulrike Domahs & Beatrice Primus (eds.), Handbuch Laut, Gebärde, Buchstabe, 337–355. (Handbooks of Linguistic Knowledge 2). Berlin/Boston: De Gruyter. https://doi.org/10.1515/9783110295993-019

Dahl, Eystein & Antonio Fábregas. 2018. Zero morphemes. In Mark Aronoff (ed.), The Oxford research encyclopedia of linguistics. Oxford: Oxford University Press. https://doi.org/10.1093/acrefore/9780199384655.013.592

Daniels, Peter T. 1991. Is a structural graphemics possible? LACUS Forum 18. 528–537.

Davidson, Andrew. 2019. Writing: The re-construction of language. Language Sciences 72. 134–149. https://doi.org/10.1016/j.langsci.2018.09.004

Dürscheid, Christa & Dimitrios Meletis. 2019. Emojis: A grapholinguistic approach. In Yannis Haralambous (ed.), Graphemics in the 21st Century, 167–183. (Grapholinguistics and Its Applications 1). Brest: Fluxus Editions. https://doi.org/10.36824/2018-graf-duer

Handel, Zev. 2019. Sinography: The borrowing and adaptation of the Chinese script. Leiden: Brill. https://doi.org/10.1163/9789004352223

Haspelmath, Martin. 2010. Comparative concepts and descriptive categories in crosslinguistic studies. Language 86(3): 663–687. https://doi.org/10.1353/lan.2010.0021

Honda, Keisuke. 2009. Homographic kanji, their ambiguity and the effectiveness of okurigana as a device for disambiguation. Written Language & Literacy 12(2): 213–236. https://doi.org/10.1075/wll.12.2.06hon

Lachmann, Thomas & Kirstin Bergström. 2023. The multiple-level framework of developmental dyslexia: the long trace from a neurodevelopmental deficit to an impaired cultural technique. Journal of Cultural Cognitive Science. https://doi.org/10.1007/s41809-023-00118-2

Linell, Per. 2005. The written language bias in linguistics. Its nature, origins and transformations. (Routledge Advances in Communication and Linguistic Theory). London: Routledge. https://doi.org/10.4324/9780203342763

Meletis, Dimitrios. 2019. The grapheme as a universal basic unit of writing. Writing Systems Research 11(1). 26–49. https://doi.org/10.1080/17586801.2019.1697412

Meletis, Dimitrios. 2020. The nature of writing. A theory of grapholinguistics. (Grapholinguistics and Its Applications 3). Brest: Fluxus Éditions. https://doi.org/10.36824/2020-meletis

Mora-Marín, David F. 2008. Full phonetic complementation, semantic classifiers, and semantic determinatives in ancient Mayan hieroglyphic writing. Ancient Mesoamerica 19. 195–213. https://doi.org/10.1017/S0956536108000345

Neef, Martin. 2005. Die Graphematik des Deutschen. (Linguistische Arbeiten 500). Berlin/Boston: De Gruyter. https://doi.org/10.1515/9783110914856

Okada, Kazuhiro. 2021. The less unit-ness of grapheme in the Japanese writing system. Paper presented at the 16th International Conference of the European Association of Japanese Studies, virtual, 28 August, 2021.

Osterkamp, Sven & Gordian Schreiber. 2021. <Th>e ubi<qu>ity of polygra<ph>y and its significan<ce> for <th>e typology of <wr>iti<ng> systems. Written Language & Literacy 24(2). 171–197. https://doi.org/wll.00052.ost

Reinken, Niklas. 2022. How can complex graphemes be identified in German? Linguistics Vanguard 8(1). 321–330. https://doi.org/10.1515_lingvan-2022-0055

Rizza, Alfredo. 2021. On 'grapheme': Recurrent problems and new reflections. In Alessia Bauer & Gaby Waxenberger (eds.), Wege zur Konfiguration der Zeichen-Phonem-Beziehung, 19–35. (ScriptandSound 3). Wiesbaden: Reichert.

Royer, Carine, Elsa Spinelli & Ludovic Ferrand. 2005. On the status of mute letters in French: Simple graphemes or part of complex graphemes? Current Psychology Letters – Behaviour, Brain & Cognition 16(2). https://doi.org/10.4000/cpl.465

Schmidt, Karsten. 2018. Phonographie und Morphographie im Deutschen. Grundzüge einer wortbasierten Graphematik. (Stauffenburg Linguistik 107). Tübingen: Stauffenburg.

Stetter, Christian. 2011. Ideographie und Alphabetschrift: Zur Ikonizität von Schriften. In Antonio Loprieno, Carsten Knigge Salis & Birgit Mersmann (eds.), Bild, Macht, Schrift: Schriftkulturen in bildkritischer Perspektive, 37–66. Weilerswist: Velbrück Wissenschaft.

Tranter, Nicolas. 2013. Logography and layering: A functional cross-linguistic analysis. Written Language & Literacy 16(1). 1–31. https://doi.org/10.1075/wll.16.1.01tra