To be cited as: Meletis, Dimitrios (2024). Alphabet. *Reference Collection in Social Sciences*. Amsterdam: Elsevier. DOI: 10.1016/B978-0-323-95504-1.00013-2.

Alphabet

Dimitrios Meletis
Department of Linguistics, University of Vienna
Sensengasse 3A, 1060 Vienna, Austria
dimitrios.meletis@outlook.com

Abstract

This article provides an overview of the alphabet as a distinctive type of writing system. It begins by tracing milestones in its historical development and spread before delving into typological considerations and offering a definition of the alphabet as a segmental phonographic system. Furthermore, the article explores several structural features characteristic of alphabetic writing systems that are claimed to have a bearing on their psycholinguistic processing, including the (in)transparency of their grapheme-phoneme correspondences, their relatively small inventory size, and their adaptability to represent a wide array of languages. Lastly, the article highlights the sociolinguistic implications of largely Eurocentric scholarly discourses on the alphabet as well as its global dissemination.

Keywords

Greek alphabet, Roman script, Latin script, segmentary, phonography, grapheme-phoneme correspondence, transparency, case distinction, bicamerality, alphabetocentrism

Key points

- The alphabet is a segmental phonographic writing system (a 'segmentary') characterized by the correspondence of individual graphemes with individual phonemes. It is distinguished from other segmentaries by its full representation of vowels and the functional and formal equivalence of consonant and vowel graphemes.
- Its historical development traces back to around 800 BCE, when the Greeks adopted the consonantal Phoenician writing system and repurposed unneeded letters to represent vowels.
- Structural features of alphabetic writing systems that influence their processing include the (in)transparency of their grapheme-phoneme correspondences, their typically small inventory size, and their versatility in representing languages.
- A feature closely associated with alphabets is the distinction between two cases (uppercase letters/majuscules and lowercase letters/minuscules).
- Alphabetocentrism, rooted in the belief of the alphabet's superiority over other types of writing systems, has influenced both scholarly discourses and international literacy initiatives.
- The global prevalence of the alphabet is primarily due to the widespread choice to adopt Roman script, partially because of its technological implementation.

1 Introduction

The alphabet is a type of writing system characterized by the correspondence of single graphic signs (graphemes) with individual phonological segments (phonemes). This distinguishes alphabets from writing systems whose graphemes correspond with moras/syllables (like Cherokee) or morphemes (such as Chinese). While the occasional broad use of the term 'alphabet' encompasses all segmental phonographic writing systems—including those that represent no or not all vowel phonemes—modern usage confines it to systems in which both consonant and vowel phonemes are written and no formal and functional distinction is made in the graphemes representing them. Etymologically, *alphabet* derives from the first two letters of the Greek alphabet, *alpha* and *beta*, which originate from the Phoenician letters *aleph* and *beth*, underlining the foundational influence of the consonantal writing system of Phoenician on the Greek alphabet.

After outlining milestones in the alphabet's historical development (Section 2), this article addresses various typological and terminological considerations (Section 3). It will then address not only the alphabet's structure but also its use as well as ideologies associated with it. This adoption of multiple analytical perspectives is particularly relevant for a comprehensive treatment of writing systems as they are often subjected to oversimplified evaluations such as "[s]ome writing systems are better than others" (Rogers, 1995, p. 31), with a tendency to favor the alphabet over other types. Such assertions overlook the complex interplay between the multiple functions assumed by writing systems, which must suit not only the languages they provide with a written form but also the physiological and cognitive makeup of human readers and writers, as well as the sociocultural and political conditions, needs, and wishes of literate linguistic communities (Meletis, 2020). To do these different perspectives justice, this article will explore both structural and psycholinguistic aspects (Section 4) as well as sociolinguistic implications (Section 5) of the alphabet.

2 Historical milestones

Around 800 BCE, the Greeks adopted the Phoenician writing system to write their own language. Phoenician had signs only for consonant phonemes, several of which were not present in the phonology of Greek. The "first Greek scribe simply did not hear [these consonants] — and thought [their] letters stood for the vowels that followed those consonants in the letter names" (Daniels, 2006, p. 10). This resulted in the first case of segmental vowel notation, making the Greek alphabet the first alphabet in the narrow sense (Healey, 1994). After it was adopted, numerous local or *epichoric* varieties emerged and came into concurrent use (Elvira Astoreca, 2021); of these, the variety of Athens—comprised of 24 letters that are still used today—was officially adopted around 400 BCE (Daniels, 2017). The Etruscans borrowed a west Greek variety around 700 BCE, and the resulting Etruscan alphabet then served as the model for the Latin alphabet. Finally, "[i]n *Roman* hands, the letters assumed the shapes of today's capital letters, and the *Roman* Church (insisting on Latin for its own use everywhere) carried it throughout Western Europe, where scribes used it for their own languages as well" (Daniels, 2017, p. 78, my emphasis). This Roman influence is the reason why, here, the term Roman (instead of the often synonymously used Latin) is employed for the script that represents the basis for numerous writing systems (see Section 3.2).

While the Roman Church insisted on unity, Eastern churches embraced diversity. This led to the creation of, among others, new scripts and alphabets for Coptic (for the last stage of ancient Egyptian), Gothic (for the extinct East Germanic language), Armenian, Georgian (for details on these four alphabets, see Gamkrelidze, 2006), as well as Glagolitic and Cyrillic. They all represent adaptations of the Greek alphabet, which served as a model in various ways, e.g., by influencing aspects such as the shapes of letters, their sound values, and their ordering.

In modern times, the alphabet is spreading mostly due to the adoption of Roman script in the process of devising writing systems for unwritten languages (Section 5).

3 Typological aspects

3.1 Alphabets as segmentaries

In a writing system, graphemes[1] correspond (mainly) with one[2] type of linguistic unit, the basic options being the phoneme, the syllable, and the morpheme. This correspondence unit commonly serves as the criterion for assuming different—necessarily idealized—types of writing systems (Joyce & Meletis, 2021): the phonographic types of segmentary (phoneme-based) and syllabary (syllable-based) as well as morpheme-based morphography, which is commonly not further divided into subtypes.

In modern approaches to writing system typology, the alphabet is regarded as only one subtype of segmentary (a term proposed by Gnanadesikan, 2017), the other subtypes being abjads and abugidas (terms coined by Peter T. Daniels by analogy with 'alphabet'; Daniels 1990). The decisive functional difference between them is their inclusion vs. omission of vowels: alphabets are fully voweled segmentaries, meaning all vowels are written. By contrast, abjads such as Arabic or Hebrew either lack vowels entirely or represent only some vowels. Finally, abugidas such as Devanāgarī or Thai prototypically represent all vowels but one. This vowel is 'inherent' to the basic consonant graphemes as it remains graphically unmarked while the remaining vowels are indicated by modifications or additions to the basic graphemes.[3] A second important difference is of formal nature: in alphabets, vowels and consonant graphemes are typically of equal size and are placed next to each other on the base line,[4] while in abjads and abugidas, potential vowel graphemes are relegated to secondary status in being smaller than and/or spatially dependent on consonant graphemes.

3.2 Alphabets and different scripts

At the typological level, the alphabet is a type of writing system. At the individual level, each alphabetic writing system (also referred to as 'alphabet') is the pairing of a specific language, e.g., English or Russian, and a given script, e.g., Roman or Cyrillic, the latter being defined here as "a set of graphic signs with prototypical forms and prototypical linguistic functions" (Weingarten, 2011, p. 16). It is necessary to make these conceptual distinctions as some of the world's scripts are used for numerous writing systems, the most obvious example being Roman script, which has been adopted for many alphabets of languages belonging to quite distinct language families (such as Indo-European, Uralic, and Austronesian). Notably, when a script is borrowed by a new language, a variety of strategies aid in adapting it to the language's needs including the dropping or repurposing of unneeded letters, the graphic alteration of existing letters, and the addition of diacritics (Daniels, 2006). At the other end of the spectrum we find scripts that are used for a single writing system, such as the Armenian script for the Armenian alphabet. In such cases, a script is, in a sense, 'custom-made' for a language, and the historical and linguistic connection between them is more direct than when a language adopts a previously existing script originally devised for a different language.

While scripts are predominantly associated with writing systems of a given type, this link is not fixed: the Hebrew and Arabic scripts, for example, while commonly used to write abjads, are also

---

[1] This refers to default graphemes, which exclude, for example, digits, special characters, and punctuation marks, which often behave differently.

[2] A notable exception is the Japanese writing system, in which syllabographic graphemes—the *kana* scripts—are used in a systematic combination with morphographic graphemes—the Chinese-derived *kanji* as well as (though more marginally) the alphabetic Roman script-*romaji*.

[3] This tripartite distinction and its inherent conflation of functional and formal criteria is, in several ways, reductive, as there is also considerable variety among the systems subsumed by these types (Gnanadesikan, 2017).

[4] An exception is Hangul, the Korean alphabet, in which the graphemes that correspond with phonemes are spatially subsegmental and arranged in segmental syllable blocks (Pae, 2024).

used for the Yiddish and Uyghur writing systems, respectively—both of which can be considered alphabetic due to their full representation of vowels.

There is also considerable variety regarding the substantiation of alphabets that transcends the prototypical material features of the written modality: to name just two examples, braille systems (which are to some degree alphabetic) function not visually but primarily via the tactile channel (Britz, 1996), and manual alphabets in diverse sign languages (Power, Grimm, & List, 2020) do not use a writing surface and—except for the hand(s)—no writing instrument.

4 Structural and psycholinguistic aspects

Despite their assignment to the same type, the world's alphabetic writing systems exhibit considerable linguistic variation.[5] One feature introducing differences is the varying transparency of correspondences between graphemes and phonemes,[6] which has occasionally even motivated the distinction of further subtypes of the alphabet, for example, between 'pure phonemic systems' such as Greek, Latin, or Finnish, and 'morphophonemic systems' such as Korean, French, or English (DeFrancis, 1989, p. 58). This is echoed in psycholinguistic research, with the influential orthographic depth hypothesis (Katz & Frost, 1992) claiming that pure phonemic (or 'shallow') writing systems are more easily learned than morphophonemic (or 'opaque') ones. Such approaches are usually restricted to the segmental level and emphasize the (psycho-)linguistic ideal of complete phonographic transparency. However, it is vital to recognize that despite their small grain size, alphabets also visualize higher-order linguistic structures such as syllables, feet, morphemes, and lexemes (for a description of the graphematic representation of such structures in German, see the chapters in Domahs & Primus 2016). Depending on the alphabet, these can prove equally if not more crucial for processes of reading and writing (Ziegler & Goswami, 2005).

A feature that is often touted as an advantage of alphabetic writing systems over non-segmental counterparts is their relatively small inventory size, ranging from 12 letters (as exemplified by the Rotokas alphabet, utilizing a subset of Roman script) to approximately thirty letters (as observed in alphabets like Russian and Georgian). This modest size stems from the direct correspondence between alphabetic graphemes and phonemes, with a language's phoneme inventory being smaller than its number of well-formed syllables or morphemes. From a psycholinguistic perspective, this feature is frequently associated with enhanced ease of learning. This warrants nuance, as inventory size represents only one aspect influencing the learning process as it may facilitate the memorization of letters. By contrast, an assumption that is under scrutiny is that the phoneme is the most accessible unit for learners acquiring literacy in the first place. As an isolated segment in what is in actuality a continuous stream of sound, it has been claimed to be an artificial epiphenomenon brought upon by segmental literacy (Davidson, 2019). This is in accordance with studies suggesting that syllabic structures may be easier to learn (see, for example, Inkelas et al., 2013), which is echoed by claims of the syllable's special role in both the phylogenetic and ontogenetic development of writing (Daniels, 2017).

Notably, the process of acquiring literacy—including processes and practices of both reading and writing—must be distinguished from the processes of using literacy once it has been acquired. Beginning readers of alphabets favor phonographic transparency, i.e., read for sound, whereas experienced readers prefer morphographic transparency, i.e., read for meaning. This appears to parallel a trend in the developmental trajectory of writing systems, wherein phonographic systems

[5] These differences are not captured by writing system typology, which is criticized by Weingarten (2011, p. 11): "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."

[6] For a proposal of how to measure the segmental transparency of alphabets quantitatively, see Neef and Balestra (2011).

tend to "evolve from being phonetically-based when they are young, towards being lexically-distinctive as they mature" (Sampson, 2018, p. 10). A notable example of this is, of course, the English alphabet, which nowadays is considered notoriously opaque with respect to letter-sound-correspondences due to its preservation of historical spellings (Daniels, 2017, p. 77–78).

Due to its segmental phonographic nature, the alphabet is often noted for its versatility in rendering languages into written form. Indeed, the International Phonetic Alphabet (IPA), the standard phonetic notation system used by linguists to describe any language, operates alphabetically and draws heavily upon symbols from the Roman script (for an analysis of IPA as an alphabet, see Neef, 2015). An often-cited counterexample of the alphabet's limited adaptability is its unsuitability for writing the varieties of Chinese (especially Mandarin): Given that morphemes in Mandarin are predominantly monosyllabic, and the language possesses a relatively restricted inventory of well-formed syllables, there exist many homophonous morphemes. When written with an alphabet, now, unlike in the morphographic system of Chinese, these morphemes would be homographs (e.g., *lù*, which could mean, among others, 'road', 'deer', or 'prosperity'), thus creating ambiguity and reducing readability.[7] Notably, alphabetic solutions do also exist for Chinese—such as the Pinyin romanization—but they are only used for special and limited contexts and functions such as digital input methods.

A feature tightly associated with the alphabetic type is the case[8] distinction between majuscules (or uppercase letters) and minuscules (or lowercase letters). Most alphabets are bicameral, meaning there are commonly two variants for representing one phoneme such as ‹B› and ‹b›, whereas most non-alphabetic writing systems are unicameral.[9] Historically, majuscules are the primary variants; the first script to develop minuscules was Roman, with a decisive moment being the emergence of the Carolingian minuscule from around 800 CE (for a systematic formal explanation of how Roman minuscules developed out of majuscules, see Brekle, 1995). The impetus behind this 'minusculization' stems from the practical utility of minuscules as they enable swifter writing and enhance readability. Subsequently, other scripts such as Cyrillic, Armenian, and Georgian[10] also adopted the concept of bicamerality. Formally, minuscules often resemble majuscule counterparts; in some instances, they are merely smaller renditions (e.g., Cyrillic ‹Ш› and ‹ш›). Functionally, minuscules are regarded as the default (Primus, 2006) as majuscules assume specific functions such as marking sentence beginnings or proper names, lending them psycholinguistic significance in reading processes. The relation between minuscules and majuscules can be conceptualized as a form of syntagmatic allography (Meletis & Dürscheid, 2022).

5 Sociolinguistic aspects

A sociolinguistic treatment of the alphabet must start with a mention of alphabetocentrism, a Eurocentric ideology that posits the superiority of the alphabet and rests upon two pillars: Firstly, it subscribes to the now-discredited teleological notion known as the *principle of unidirectional development*, which assumes that all writing systems progress through distinct typological stages, culminating inevitably in the alphabetic form (Gelb, 1963). Secondly, some scholars claim that alphabetic literacy yields profound cognitive 'consequences' and ascribe to it significant social, political, and economic benefits, effectively arguing for its pivotal role in shaping the trajectory of

[7] This situation is only marginally alleviated when lexical tone is represented by diacritics.

[8] The term 'case' comes from the days of metal type printing, where capital letters were stored in the upper compartments of a printer's case ('uppercase') while the smaller letters were kept in the lower compartments ('lowercase').

[9] Exceptions include the Warang Citi abugida and the Cherokee syllabary, which both have majuscules and minuscules.

[10] The Georgian alphabet now uses letters of the Mkhedruli variant, which can be considered 'lowercase', while the 'uppercase' letters—called Mkhedruli Mtavruli—are only used rarely, so that Georgian can de facto be considered unicameral.

Western civilization (Logan, 2004; see also *Literacy*). This techno-deterministic view entails the marginalization of alternative writing systems and literacies, devalorizing them as inferior or 'restricted'. A poignant example is the historical scrutiny directed towards Chinese morphography, echoed in unsuccessful 20th century attempts to Romanize the writing system. This bias has further manifested in the early linguistic study of writing systems, marked by a tendency towards exoticization and inadequate analysis of non-alphabetic scripts, while psychological research on reading has disproportionately favored alphabetic systems, and particularly the non-prototypical alphabet of English, which has resulted in models of reading being skewed (Share, 2014).

The pervasive scholarly attention accorded to the alphabet is reflected in its global prevalence, particularly evident in the myriad alphabets based on Roman script. The fact that it is the most widespread script can be attributed to several factors: On the one hand, it is the script that Western linguists involved in devising new writing systems, who include missionaries and scholars affiliated with the *Summer Institute of Linguistics* (SIL), have traditionally been best acquainted with (Coulmas, 1996). On the other hand, in a globalized world marked by fast-spreading and evolving digital communication, language communities often seek participation in 'developments on a larger scale' (Unseth, 2005) and thus gravitate towards scripts with established technological infrastructure, the basis of which is their encoding in Unicode. This trend is exemplified by recent adoptions of the Roman script in literacy initiatives across Africa (Roberts, 2023).

Conclusion

Typologically, the alphabet is characterized by its segmental phonographic nature, i.e., the correspondence of graphemes to individual consonant and vowel phonemes. Emerging historically from the Phoenician abjad and developing through the Greek and Latin scripts, the alphabet has spread globally, becoming the basis for the writing systems of numerous diverse languages. However, the teleological notion of the alphabet as an inevitable—and ideal—end stage of the history of writing reflects a Eurocentric bias that results in an idealization of the alphabet as well as the marginalization of other writing systems. While alphabets do exhibit several frequently commended features such as small inventory sizes and a certain versatility in representing diverse languages, it is essential to recognize that both the academic study of writing systems as well as the real-life choice of a script and (type of) writing system employed by a literate community encompass not only structural and psycholinguistic aspects but also sociolinguistic considerations. This underlines the complex interplay between writing systems, languages, and cultures and the fact that the alphabet is but one of its many—and manifold—products.

References

Brekle H E (1995). Neues über Groß- und Kleinbuchstaben. Theoretische Begründung der Entwicklung der römischen Majuskelformen zur Minuskelschrift. *Linguistische Berichte 155*, 3–21.

Britz K (1996). Blindenschrift (Braille). In Günther H & Luwdig O (eds.) *Schrift und Schriftlichkeit/Writing and its use*. Berlin/Boston: De Gruyter. 1617–1623.

Coulmas F (1996). *The Blackwell encyclopedia of writing systems*. Oxford: Wiley-Blackwell.

Daniels P T (1990). Fundamentals of grammatology. *Journal of the American Oriental Society 110(4)*, 727–731.

Daniels P T (2006). On beyond alphabets. Written Language & Literacy *9(1)*, 7–24.

Daniels P T (2017). Writing systems. In Aronoff M & Rees-Miller J (eds.) *The handbook of linguistics* (2nd edn.). Oxford/Malden: Wiley-Blackwell. 75–94.

Davidson A (2019). Writing: The re-construction of language. *Language Sciences 72*, 134–149.

DeFrancis J (1989). *Visible speech: The diverse oneness of writing systems*. Honolulu: University of Hawaii Press.

Domahs U & Primus B (eds.) (2016). *Handbuch Laut, Gebärde, Buchstabe*. Berlin/Boston: De Gruyter.

Elvira Astoreca N (2021). *Early Greek Alphabetic Writing. A Linguistic Approach*. Oxford, Philadelphia: Oxbow Books.

Gamkrelidze T V (2006): Europe, Christian: Alphabets. In Brown K (ed.) *Encyclopedia of language and linguistics* (2nd edn.). Amsterdam: Elsevier. 295–305.

Gelb, I J (1963). *A study of writing* (2nd edn.). Chicago: University of Chicago Press.

Gnanadesikan A E (2017). Towards a typology of phonemic scripts. *Writing Systems Research 9(1)*, 14–35.

Healey J F (1994). Alphabet: Development. In Asher R E (ed.) *The encyclopedia of language and linguistics*. Oxford: Pergamon Press. 74–79.

Inkelas S, Johnson K, Lee C, Minas E, Mulcaire G, Yong Keng G & Yuasa T (2013). Testing the learnability of writing systems. *Annual Meeting of the Berkeley Linguistics Society 39*, 75–89.

Joyce T & Meletis D (2021). Alternative criteria for writing system typology. Cross-linguistic observations from the German and Japanese writing systems. *Zeitschrift für Sprachwissenschaft 40(3)*, 257–277.

Katz L & Frost R (1992). The reading process is different for different orthographies: The orthographic depth hypothesis. In Frost R & Katz L (eds.) *Orthography, phonology, morphology, and meaning*. Amsterdam: Elsevier. 67–84.

Logan R K (2004). *The alphabet effect: A media ecology understanding of the making of Western civilization* (2nd edn.). New York: Hampton Press.

Meletis D & Dürscheid C (2022). *Writing systems and their use. An overview of grapholinguistics.* Berlin/Boston, De Gruyter Mouton.

Meletis D (2020). *The nature of writing. A theory of grapholinguistics*. Brest: Fluxus Editions.

Neef M & M Balestra (2011). Measuring graphematic transparency – German and Italian compared. *Written Language & Literacy 14(1)*, 109–142.

Neef M (2015). Writing systems as modular objects: Proposals for theory design in grapholinguistics. *Open Linguistics 1*, 708–721.

Pae, H (2024). *Analyzing the Korean alphabet. The science of Hangul*. Cham: Springer.

Power J M, Grimm G W, and List J-M (2020). Evolutionary dynamics in the dispersal of sign languages. *Royal Society Open Science 7*, 191100.

Primus B (2006). Buchstabenkomponenten und ihre Grammatik. In Bredel U & Günther H (eds.) *Orthographietheorie und Rechtschreibunterricht*. Tübingen: Niemeyer. 5–43.

Roberts D (2023). Roman script orthography development in Africa: Historical and contemporary perspectives. In Malatesha Joshi R, McBride C A, Kaani B & Elbeheri G (eds.) *Handbook of literacy in Africa*. New York: Springer Nature. 59–78.

Rogers H (1995). Optimal orthographies. In Taylor I & Olson D R (eds.) *Scripts and literacy: reading and learning to read alphabets, syllabaries and characters*. Dordrecht: Springer. 31–43.

Sampson G (2018). From phonemic spelling to distinctive spelling. *Written Language & Literacy 21(1)*, 3–25.

Share D L (2014). Alphabetism in reading science. *Frontiers in Psychology 5*, 752.

Unseth P (2005). Sociolinguistic parallels between choosing scripts and languages. *Written Language & Literacy 8(1)*, 19–42.

Weingarten R (2011). Comparative graphematics. *Written Language & Literacy 14(1)*, 12–38.

Ziegler J & Goswami U (2005): Reading acquisition, developmental dyslexia, and skilled reading across languages: a psycholinguistic grain size theory. *Psychological Bulletin 131(1)*, 3–29.