

# Orthography

GINA COOKE

Illinois State University, USA

“Languages get the writing systems they deserve.”

M.A.K. Halliday

The study of written language involves multiple disciplines. Cognitive psychology and neuroscience contribute theoretical and developmental models of literacy in different languages. Historians rely on written records to illuminate the past, and some study the origins and evolution of writing itself. Linguistic anthropologists may research the intersection of writing and community, including non-standard writing practices such as graffiti or texting. But for more than a century, linguists and anthropologists have targeted speech as their primary focus, largely discounting writing as contingent upon speech. At the inception of the modern field, Ferdinand de Saussure charged his predecessors with “confus[ing] language and writing” ([1916] 2011, 24), and set his focus on the primacy of speech. Structural linguistics doubled down on this differentiation, characterizing writing as parasitic to speech. This division between speech and writing was critical in the early years of modern linguistics, when indigenous languages were largely unwritten but no less worthy of scientific study than languages with long literary traditions. Twentieth-century linguistics sought to differentiate speech, which is innate human behavior and far older than writing, from written language, which is neither innate nor universal. “Still today,” Saussure complained, “intelligent men confuse language and writing” ([1916] 2011, 24), and his lament often remains the case in the twenty-first century.

While the study of written language crosses disciplinary barriers, the subject of orthography in particular often ends up in an academic no-man’s land, reserved for prescriptive grammarians, proofreaders, and orthographic reformers. This entry offers an overview of what an orthography is and how it differs from an *alphabet*, a *writing system*, and a *script*, and illuminates how orthographies work structurally and within communities.

## Definitions

Academic studies of orthography have traditionally been concerned with the prescriptive teaching and correct performance of literacy, rather than the descriptive study and analysis of writing systems. The past three decades, however, have seen a surge in the study of orthography, from orthographic depth and literacy development to sociological and anthropological studies. Scholars who study orthography differentiate the term from other similar nomenclature: *orthography* is sometimes used interchangeably with

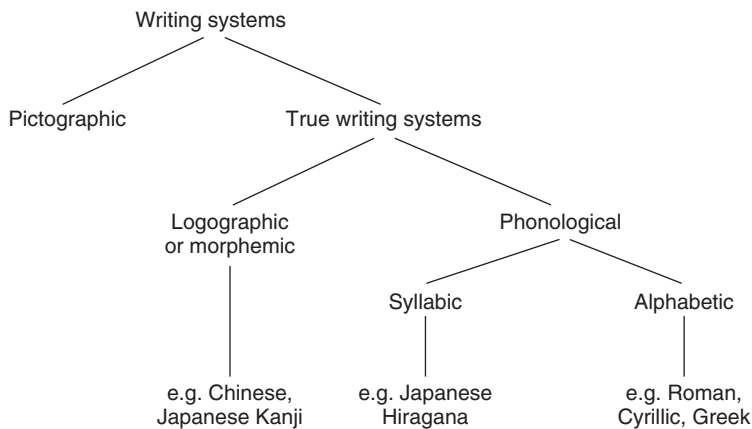
*The International Encyclopedia of Linguistic Anthropology*. Edited by James Stanlaw.

© 2021 John Wiley & Sons, Inc. Published 2021 by John Wiley & Sons, Inc.

DOI: 10.1002/9781118786093.iela0297

*writing system* as well as with *spelling*, along with *alphabet* and *script*. In general conversations, such connotations are harmless; in linguistic theory and analysis, however, some distinctions are called for.

A true *writing system*, as opposed to pictographic symbols, matches symbols to linguistic structures rather than to ideas or concepts. Writing systems are graphical means for recording language, and they can relate to speech according to varying principles: characters may represent words, morphemes, syllables, consonants, or vowels; they may also mark phonetic features like tone, stress, palatalization, and even silence. This variation among linguistic structures represented in writing is customarily divided into *logographic* and *phonological* writing systems (see Figure 1). Phonological systems are further divided into syllabic and segmental (or phonemic) systems, which usually have a character inventory of *alphabetic letters*. Segmental systems are further divided into fully segmental systems representing both vowels and consonants; those that represent only consonants, or *abjad* systems; and alphasyllabary or *abugida* systems whose characters represent consonant-vowel sequences as units (Daniels and Bright 1996, xxxix). An *orthography* is a concrete, language-specific instantiation of any type of writing system, or the “relationship between a script and its language” (Katz and Frost 1992, 147). An orthography consists of a language’s proprietary inventory of written symbols and the conventions for using the symbols. This includes *spelling*, capitalization, punctuation, word form, and word breaks. Russian, Arabic, and Hmong all have discrete, unique, conventional *orthographies* that are examples of segmental writing systems, and each of them uses a different *script*. Russian uses the *Cyrillic script*, as do Bulgarian and Ukrainian, but each language has its own specific *alphabet*, or inventory of letters. Arabic, Urdu, and Farsi all use the Perso-Arabic *script*, but each of them has its own *alphabet* and *orthography*. Likewise, the Hmong Romanized Popular Alphabet (RPA) uses a Latin script like English and Vietnamese do, but Vietnamese has 29 letters and diacritics, English uses 26 letters with no diacritics, and Hmong uses several consonant letters to represent tones rather than consonant segments. Some languages have more than one orthography, often called digraphia (DeFrancis 1984), although a language



**Figure 1** Traditional taxonomy of writing systems. Source: Stubbs (1980, 48). Reproduced by permission of Taylor & Francis.

may also have more than two writing systems as well, including the use of more than one script. Hmong has been recorded with more than a dozen orthographic approaches using multiple scripts (Chinese logograms, Thai syllabary, Latin script, and an invented script, Pahawh Hmong), and Japanese uses both *kanji* logograms and *kana* syllabary. Having different or even competing orthographies may reflect different linguistic structures, like the representation of Hindustani languages in both Perso-Arabic segmental and Devanagari abugida scripts. However, different orthographies can also reflect different linguistic, historical, and cultural identities.

## Orthographic linguistics

A close look at real-life orthographies proves to be less orderly than the abstract, conceptual division of writing systems into meaning-versus-sound-based frameworks (Figure 1). As indicated by the presence of alphasyllabaries among phonological systems, and the frequent characterization of Chinese as a logosyllabary, such discrete divisions are more of an ideal conceptualization. Orthographic linguists stress that real orthographies are largely *mixed systems*, representing both phonological and lexical or morphological information, regardless of their specific mechanics or where they fall in the traditional analysis. Many alphabetic systems preserve morphological information and identity, often at the expense of graphophonemic consistency. Alphabetic orthographies can be further classified according to their *orthographic depth* (Lieberman et al. 1980): a *shallow* or *transparent* orthography bears a high degree of graphophonemic isomorphy, while graphemes in a *deep* or *opaque* system can have multiple pronunciations, including silence, and segments can have multiple spellings. “Shallow orthographies are characteristic of languages in which morphemic relatives have consistent pronunciations” (Katz and Frost 1992, 150), so deeper orthographies like English or French are driven by the need to maintain morphological identity in lexical families like *sign~signal~designate*, *heal~health*, *blond~blonde*, or *collègue~collégial*, despite changes in pronunciation. This is to be expected, however, because in any segmental system, regardless of depth, phoneme–grapheme correspondences are delimited and constrained by morphemic boundaries. Even the highly transparent Finnish orthography preserves its agglutinative morphology.

While a purely segmental system would theoretically offer a one-to-one correspondence between phonemes and graphemes, this is “an abstract principle that is at best approximated by actual writing systems” (Coulmas 2003, 18). Even alphabetic orthographies have mechanisms for including *lexical* or *morphological* information as well as segmental. In addition to preserving morphological identity, segmental spellings may be etymologically driven, like a <ph> in Latin words of Greek origin, or preserve foreign spellings, like Sanskrit graphemes in Tatsama loan-words. Furthermore, most segmental systems have affordances for differentiating distinct words; the most transparent orthographies have a high degree of homonymy and thus rely on context, as in Finnish *kola* (cola, plow) or Turkish *ocak* (January, oven). Even highly transparent systems may have graphemic options for differentiating homophones, as Spanish *hoya* (pit) and *olla* (cooking pot), or German *bis* (to) and *Biss*

(bite). The Korean Hangul script was invented in the fifteenth century to replace the ill-suited Chinese Hanzi characters; nonetheless, Hangul still uses Chinese characters to semantically differentiate homophones.

Not only is the traditional analysis of writing systems in Figure 1 artificial; some orthographic linguists also suggest that the divide between phonological and logographic systems is Western-centric. Labels like *semantic* or *ideographic* “are popularly attached to Chinese characters by Western writers,” says DeFrancis, “but this should not obscure the fact that phoneticity ... far surpasses iconicity, which actually approaches zero” (1989, 114). All writing systems encompass both pronunciation and meaning according to various principles, and the two are not mutually exclusive. This is because “spelling’s job is making sense” (Cooke 2012). As Coulmas explains, “[t]he communication of *meaning* is the primary purpose of most writing” (2003, 18, emphasis added); Ram Frost concurs: “writing systems are primarily designed to represent *meaning* to readers” (2012, 274, emphasis added). In fact, the phonological structure represented by a written language, that is, its *orthographic phonology*, is not separate from the language’s meaning, but rather one of the aspects of written language that serves to specify and differentiate meaning, along with morphology, syntax, and sometimes etymology. Although some scholars use the terms *semantic* and *phonetic* to describe opposite ends of the orthographic spectrum, those terms are not optimal for discussing real orthographies. First, semantics are a consideration for *any* working spelling system, so a discussion of an orthography’s *morphological*, *lexical*, or *syntactic* mechanisms is more accurate. Second, the sound structure and information conveyed by an orthography is *phonological*, in service of clarifying and differentiating psychological meaning, not *phonetic*, which refers to the precise physical aspects of linguistic sound. Conventional orthographies generally do not capture *phonetic* differences, like regional dialects or foreign accents; they record only those aspects of pronunciation necessary to convey meaning to those who know and use the language. Indeed, “the discrete nature of alphabetic notation makes the phonemes of a language appear a more clearly defined set than they actually are” (Coulmas 1996, 404).

## Orthography and culture

In addition to linguistic structures, non-linguistic factors can also have an effect on orthographies. Serious study of orthography requires attention to the history and evolution of a given system, as well as to the communities of people who use it, how said communities identify themselves, and the conflicts that can arise around orthographic differences, standards, and reforms. Many a scholar of spoken tongues has joked that the difference between a language and a dialect is an army or a flag; likewise, when it comes to written language, it is not always easy to identify where one system ends and another begins. Portuguese and Brazilian orthographies have a small set of minor differences, for example, similar to UK and US English. More fraught examples include

writing the same or very similar languages in different scripts, like Hindi's Devanagari script and Urdu's Perso-Arabic system, or Serbian's Cyrillic writing and Croatian's Latin script; this often occurs in geopolitical areas marked by military and social conflict.

While some writing systems have been invented for specific languages, generally, no one person or committee directs the diachronic development of an orthography. Some linguistic communities have orthographic standard-bearing institutions, like L'Académie Française or the Korean Ministry of Culture, but most do not. Some orthographic communities undergo deliberate orthographic reforms with wildly variable levels of success, however defined. Some reforms are minor, like Greek's monotonic reforms of 1982, while others are major and have historic implications, like Atatürk's 1928 replacement of the Ottoman Turkish Perso-Arabic alphabet with a Latin-based alphabet, one of a series of sweeping cultural changes. The twentieth century witnessed orthographic reforms in German, Portuguese, Czech, and Chinese, among others, each with intertwined historical, sociopolitical, and economic concerns that are still debated.

In the same way that early twentieth-century linguists asserted that all spoken language merited study regardless of the presence of a writing system, it is important that modern scholars assert the merits of studying written language, for orthography makes human thought visible as text.

SEE ALSO: Literacy; Reading; Writing and Writing Systems: Introduction; Writing and Writing Systems: Classification of Scripts; Writing and Writing Systems: History; Writing and Writing Systems: Sociolinguistic Aspects

## REFERENCES AND FURTHER READING

---

- Bloomfield, Leonard. (1933) 1984. *Language*. Chicago: University of Chicago Press.
- Chomsky, Carol. 1970. "Reading, Writing, and Phonology." *Harvard Educational Review* 40 (2): 287–309. <https://doi.org/10.17763/haer.40.2.y7u0242x76w05624>.
- Cooke, Gina. 2012. "Making Sense of Spelling – Gina Cooke" (video file). TED Ed (website), 25 September. <https://ed.ted.com/lessons/making-sense-of-spelling-gina-cooke> (accessed 1 May 2020).
- Coulmas, Florian. 1996. *The Blackwell Encyclopedia of Writing Systems*. Oxford: Blackwell. <https://doi.org/10.1002/9781118932667>.
- Coulmas, Florian. 2003. *Writing Systems: An Introduction to Their Linguistic Analysis*. Cambridge: Cambridge University Press.
- Crystal, David. 2012. *Spell It Out: The Singular Story of English Spelling*. London: Profile Books.
- Daniels, Peter T., and William Bright, eds. 1996. *The World's Writing Systems*. New York: Oxford University Press.
- DeFrancis, John. 1984. "Digraphia." *Word* 35 (1): 59–66.
- DeFrancis, John. 1989. *Visible Speech: The Diverse Oneness of Writing Systems*. Honolulu: University of Hawai'i Press.
- Frost, Ram. 2012. "Towards a Universal Model of Reading." *Behavioral and Brain Sciences* 35 (5): 263–279. <https://doi.org/10.1017/S0140525X11001841>.
- Horobin, Simon. 2013. *Does Spelling Matter?* Oxford: Oxford University Press.

- Joshi, R. Malatesha, and P.G. Aaron, eds. 2006. *Handbook of Orthography and Literacy*. Mahwah, NJ: Lawrence Erlbaum.
- Katz, Leonard, and Ram Frost, eds. 1992. *Orthography, Phonology, Morphology, and Meaning*. Advances in Psychology, 94. Amsterdam, the Netherlands: Elsevier.
- Lieberman, Isabelle Y., Alvin M. Liberman, Ignatius Mattingly, and Donald Shankweiler. 1980. "Orthography and the Beginning Reader." In *Orthography, Reading and Dyslexia*, edited by James T. Kavanagh and Richard L. Venezky, 137–153. Baltimore, MD: University Park Press.
- Saussure, Ferdinand de. (1916) 2011. *Course in General Linguistics*. Translated by Wade Baskin, edited by Perry Meisel and Haun Saussy. New York: Columbia University Press.
- Sebba, Mark. 2007. *Spelling and Society*. Cambridge: Cambridge University Press.
- Stubbs, Michael. 1980. *Language and Literacy: The Sociolinguistics of Reading and Writing*. London: Routledge & Kegan Paul.
- Upward, Christopher, and George Davidson. 2011. *The History of English Spelling*. Chichester: Wiley Blackwell.
- Venezky, Richard L. 1999. *The American Way of Spelling: The Structure and Origins of American English Orthography*. New York: Guilford Press.