John Benjamins Publishing Company

Typology of writing systems

Special issue introduction

Terry Joyce & Susanne R. Borgwaldt

Tama University, Japan/TU Braunschweig, Germany

1. Sixth International Workshop of the Association of Written Language and Literacy (AWLL) in Braunschweig, Germany

The International Workshop on Writing Systems "Typology of Writing Systems", the sixth workshop in the biennial series organized by the Association of Written Language and Literacy, was held in Braunschweig, Germany, on September 18/19, 2008.

Rebecca Treiman, the Burke and Elizabeth High Baker Professor of Child Developmental Psychology at Washington University in St. Louis, presented the invited key note talk; a paper co-authored with Brett Kessler entitled "Similarities among the shapes of writing and their effects on learning". During the workshop, 19 talks and 3 posters were given that informed the participants about various aspects of typologically diverse writing systems, covering African scripts, Bengali, Danish, Hebrew, Japanese and Korean, amongst others. Most of the contributions to this special issue on "Typology of Writing Systems" have developed out of papers delivered during the workshop.

2. Brief review of research relating to the typology of writing systems

Against the broader background of general neglect, until comparatively recently, of written language in general and writing systems more specifically as topics of serious linguistic research, it is perhaps not so surprising that historically there have been relatively few proposals for typologies of writing systems. This section presents a short review of some of the most influential and controversial in order to identify and highlight some of the issues and themes that have shaped approaches to writing system typologies (also see Coulmas (1996b) for discussion of typologies of writing systems and their objectives).

One of the earliest classifications of writing systems was proposed by Taylor (1883) (as cited in Daniels, 1996a, 2001; DeFrancis, 1989; Diringer, 1962; Hill, 1967; Trigger, 2004). DeFrancis (1989: 59), for instance, describes Taylor's classification as being "an evolutionary model consisting of a succession of five stages", from (1) pictures, (2) pictorial symbols, (3) verbal signs, (4) syllabic signs, to (5) alphabetic signs, with the first three stages being referred to as ideograms and the last two stages as phonograms. The influence of this early classification may be detected in a number of subsequent typologies that have grappled with conceptualizing the essence of the distinction between phonographic and nonphonographic writing, where the variety of proposed alternative terms would seem to testify to its elusive nature.

Unquestionably, the seminal work on writing systems is Gelb (1952) which attempted to lay the foundations for the scientific study of writing. Although Gelb excluded from his classification pictures as a form of writing, (1) pictorial representation and (2) mnemonic devices were seen as forerunners of writing. Those two categories are, however, distinguished from the important category of full writing, which includes the three sub-categories of (3) word-syllabic, (4) syllabic, and (5) alphabetic. Despite this insightful distinction, Gelb's classification is undoubtedly flawed by his zeal to present a teleological account of writing, placing the alphabet as the final stage in an evolution via logography and syllabary (see Coulmas (1996a); Daniels (1990, 2001); Rogers (2005); Sproat (2000), and Trigger (2004) for discussion of problems with Gelb's classification).

Another early influential typology is that proposed by Diringer (1962), which, as Hill (1967) observed, is very similar to that proposed by Taylor (1883). Like Gelb's (1952) classification, Diringer's classification is also greatly influenced by the view that the alphabet represents the "most flexible and useful method of writing even invented" (Diringer 1962: 24). Although Diringer also makes a distinction between what he calls *embryo-writing* and *full writing*, his notion of full writing is more inclusive than Gelb's. Thus, under full writing, Diringer includes the five categories of (1) pictography, (2) ideography, (3) analytic transitional scripts, (4) phonetic scripts, and (5) alphabetic writing.[1] In outlining his own classification, Hill (1967) directed three criticisms towards Diringer's (1962) classification. The first point is that phonetic scripts should include both syllabaries and alphabets. The second point is that alphabetic scripts can be used in various ways, while the third criticism relates to the term ideographic. Hill claimed that his classification deals with these points and places "every system of writing in relation to that which all systems represent, language" (Hill 1967: 92), with the classification consisting of three divisions – discourse systems (with the caveat that these are only partial systems), morphemic systems, and phonetic systems.

As an early attempt to move away from a historically-orientated approach, the classification proposed by Haas (1976) is more conceptual in nature, being based on a set of three binary choices. The first choice is *derived-original*; pictographs are regarded as original because they do not correspond to speech in a regular way. The second choice is *empty-informed*; whether or not a graphic unit directly determines a meaning. The third choice is *motivated-arbitrary*; whether or not the relation between graphic unit and referent is pictorial. These choices are logically independent, but not all of the combinatory possibilities are real. This is because an empty script, for example, cannot also be motivated, so this scheme actually only recognizes five kinds of script. However, while these contrasts are useful in differentiating types of pictorial representation, there is essentially only one contrast between other scripts: the key contrast of empty-informed. Haas (1976, 1983) refers to scripts distinguished by this contrast as being either *cenemic* or *pleremic*. From the Greek word κενός meaning 'empty,' in a cenemic writing system, the graphic units only represent sounds and are, therefore, empty of semantic reference, such as alphabets and syllabaries. In contrast, from the Greek πλήρης meaning 'full', pleremic refers to writing systems where the graphic units are semantically informed, denoting both sounds and meanings, such as Japanese kanji.

The next classification that must be singled out for mention is that of Sampson (1985); not least for the considerable debate that it has inspired (see, for example, DeFrancis (1989, 2002), DeFrancis & Unger (1994), Sampson (1994), and Unger & DeFrancis (1995)). In Sampson's classification, the first distinction made is between *semasiographic* and *glottographic* writing systems.[2] At the next level, *glottographic* is divided into *logographic* and *phonographic*. Under logographic, Sampson (1985) provides for what he considers to be a logical possibility, the polymorphemic unit, although he acknowledges that no systems based on polymorphemic units actually exist. Although the non-existence of systems based on polymorphemic units would seem to render the term logographic redundant, Sampson uses it to refer to Chinese characters. Another aspect of Sampson's classification that has prompted debate was the inclusion of a featural type consisting solely of Korean Hangul.

The second major work of the 1980s on writing systems is DeFrancis' (1989) widely-cited book, which presented his writing classification scheme. At the heart of DeFrancis' classification is the dichotomy between what he refers to as partial and full writing systems, and directly linked to that, DeFrancis' conviction in the phonetic basis of all full writing systems. In line with his belief that writing is simply the visual representation of speech, DeFrancis' (1989) scheme distinguishes between six types of systems: (1) 'pure' syllabic systems (including Linear B, kana, and Cherokee); (2) morpho-syllabic systems (including Sumerian, Chinese, and Mayan); (3) morpho-consonantal systems (with Egyptian); (4) 'pure' consonantal

systems (including Phoenician, Hebrew, and Arabic); (5) 'pure' phonemic systems (including Greek, Latin, and Finnish); and (6) morpho-phonemic systems (including English, French, and Korean).

As a positive sign of the steadily growing interest in writing systems, the early 1990s witnessed the appearance of a handful of typology proposals. For instance, Daniels (1990, see also 1996 and 2001) has argued for the recognition of two other script types; abjads and abugidas,[3] as a solution to inadequacies with traditional tripartite classifications (e.g. Gelb's (1952) classification of word-syllabic, syllabaries, and alphabets), and subsequently suggested that there are six fundamentally different kinds of writing systems (Daniels 1996, 2001). The six types are (1) logosyllabary (morphosyllabary), (2) syllabary, (3) abjad (Semitic-type script), where each character stands for a consonant, (4) alphabet (Greek-type script), (5) abugida (Sanskrit-type script), where each character stands for a consonant accompanied by a particular vowel, with other vowels indicated by additions to the consonant symbol, and (6) featural, where the shapes of the characters correlate with phonetic features of designated segments. Daniels (2001: 68) claims that "once abugidas are distinguished from syllabaries, a different historical sequence can be identified, which no longer privileges the alphabet teleologically". A little later, Faber (1992) proposed a typology that distinguishes five categories based on the dimensions of (1) logographic vs. phonographic, (2) syllabically linear vs. segmentally linear, (3) complete vs. defective, and (4) syllabically encoded vs. segmentally encoded. In line with her claim that, rather than being a necessary precursor, segmentation ability is a consequence of alphabetic writing, Faber's typology yields a narrow definition of the alphabet as a segmentally linear, complete orthography, although Chinese characters are classified under logographic. Another classification of the period is that of Coulmas (1992) which draws on Haas' (1976) distinction of pleremic and cenemic writing systems in developing a classification of seven general types. Under the first division of pleremic systems, Coulmas includes (1) logograms + phonograms (e.g. Hittite hieroglyphs), (2) logograms + phonograms + determinatives (e.g. cuneiform), and (3) morphosyllabic signs (e.g. Chinese). Under the second category of cenemic writing systems come (4) syllabary (e.g. Japanese kana), (5) consonantal alphabet (e.g. Phoenician), (6) alphabet (e.g. Roman) and (7) alphabet with independent vowel letters and integrated consonant-vowel letters (e.g. Ethiopic).

While not attempting a classification of writing systems on the scale of DeFrancis (1989), in a journal article (DeFrancis & Unger 1994) and related book chapter (Unger & DeFrancis 1995), DeFrancis and Unger have argued for what they term a 'realistic' view of writing system typology. Essentially, their view focuses on the theoretical continuum between 'pure phonography' and 'pure logography'. In contrast to their claims that 'naïve' typologies assume two distinct groupings

of writing systems falling towards the two extremes with an empty middle space, DeFrancis and Unger posit the actual range of writing systems within the middle area of the continuum. Of the six writing systems marked within this middle area, while Finnish is located furthest towards the pure phonography extreme (with French and English progressively closer to the center), Chinese is positioned furthest towards the pure logography side, with Japanese more central than Chinese (and Korean more central still).

The final two typologies of writing systems that must be included within this short review are those of Sproat (2000) and Rogers (2005) which share a similar approach. The break from the conventional tree-format frequently employed in earlier classifications, emerging with DeFrancis and Unger's (1994) continuum, is taken a step further by Sproat (2000) who arranges writing systems according to two-dimensions: the type of phonography and amount of logography involved in a system. Sproat's definition of logography is perhaps more inclusive than traditional definitions, for he regards "any component of a writing system as having a logographic function if it formally encodes a portion of nonphonological linguistic structure, whether it be a whole morpheme or merely some semantic portion of that morpheme" (2000: 134). Although Rogers (2005) has more recently adopted Sproat's basic approach of two organizing dimensions, regarding it as an improvement over the earlier tree-based classifications, there are two important differences in Rogers' classification. The first difference is in the types of phonography that Rogers recognizes. Claiming that Sproat's dimension of phonography is rather unstructured, Rogers distinguishes between abjad, alphabetic, abugida, moraic, and syllabic under types of phonography. The other major difference is the label for the second dimension, where Rogers opts for the term 'amount of morphography' (which he seeks to differentiate from the related but separate notion of orthographic depth).

While it is beyond the scope of this short review to acknowledge all the perspectives and debates that have influenced the development of writing systems typologies, a few recurring issues may be discerned. One enduring concern has been to meaningfully characterize the distinction between phonographic, or cenemic, writing systems from non-phonographic, or pleremic, writing systems, where the frequently-impassioned debate has reflected divergent views about writing and its relationship to speech and language. A second challenge that has motivated some of the classifications has been to identify and characterize the meaningful categories within the range of graphemic conventions witnessed across all writing systems. As Coulmas (1996b) observes, typologies need to strike an appropriate balance between including too many types that may obscure important commonalities and only recognizing too few types that may mask deep insights about writing systems. The more recent recognition accorded to abjads

and abugidas, for instance, may be regarded as a positive improvement over earlier typologies with fewer categories, especially those that sought to portray the alphabet as a teleological achievement. Finally, although the notion of orthographic depth (Katz & Frost 1992), which was formulated primarily to account for varying degrees of consistency in grapheme-phoneme correspondences within the context of investigating the psychological processes of reading, may arguably be linked to the almost axiomatic assumption within typologies that there are no 'pure' writing systems (which can be traced back to Gelb (1952)), to the extent that the most recent typologies of Sproat (2000) and Rogers (2005) attempt to address the broad phenomenon, it reminds us how typologies of writing systems should be as informative as possible about the implications of how writing systems differ for understanding the cognitive processes involved in writing and reading.

3.   Special issue contributions to typology of writing systems research

In this section, we introduce the papers included within the special issue and seek to briefly comment on their contributions to the theme of typology of writing systems. Before turning to the individual papers, at this point, we would acknowledge that none of the present contributions attempts to propose comprehensive typologies that endeavor to cover all writing systems. However, we firmly believe that all the papers deserve careful attention for the discussions that they offer on a number of key linguistic matters that directly impact on research issues surrounding the typology of writing systems. As Coulmas (1996b) astutely notes within his review of typologies and their objectives, typologies of writing systems draw directly on theoretical notions of linguistic analysis. Elaborating further on the interaction, Coulmas (1996b: 1387) also writes as follows:

> "Since writing represents language, typologies of writing systems that are based on the units and processes by means of which this is accomplished can deepen our understanding of language, while a sharpening of the notions for analyzing the units of language can help to improve such typologies."

We see the papers of this special issue as contributing directly to the theme of typology of writing systems in the spirit of focusing on and 'sharpening' various important issues that are essential for the advancement of typological research.

The first paper by Rüdiger Weingarten particularly exemplifies this spirit with his proposal for comparative graphematics as a linguistic framework for developing meaningful comparisons of the world's writing systems and for advancing typological research. After positioning comparative graphematics as a subfield of comparative linguistics, Weingarten's introduction maps out much of the scope

and implications of comparative graphematics for a number of core topics for writing system research, including the utilization of the comparative approach for decipherment, the application to research on the creation and reform of writing systems, and the issue of an optimal orthography, as well as more cognitive concerns relating to models of reading and writing, learnability comparisons for writing systems, and the consequences of bi-literacy. As an informative illustration of comparative graphematics, the main focus of Weingarten's paper is to present a detailed discussion of the relation between the gemination of consonant letters and the graphemic representation of long consonants with examples from numerous writing systems. Interestingly, Weingarten's careful analysis highlights how graphemic constructions may undergo shifts in their function.

The interesting contribution to writing systems research of the second paper by Rebecca Treiman and Brett Kessler undoubtedly lies in its fairly unique focus on letter-shape similarities and their effects on learning. As they stress, although writing systems are usually compared in terms of different representational levels, little attention has been devoted to the shapes of the Latin letters and the levels of similarity within the set. Their paper carefully details a series of new analyses conducted on existing data about children copying and printing Latin letters. The consistent pattern in their results indicates that young children generally perform better at copying and writing the more common b-type letters (where the vertical stem, or hasta, is followed by appendage, or coda, to the right) than d-type letters (with the less common, reversed coda-hasta arrangement). As Treiman and Kessler argue, their results suggest that implicit awareness of shape frequencies influences children's early learning of letter shapes. These findings relating to shape similarities would seem to have interesting implications for studies of the early stages of learning to write with other writing systems and for comparative studies into the learnability of different writing systems.

Of some relevance to the rather thorny issue for many typologies of writing systems relating to how to portray the fundamental distinction between cenemic, or phonographic, writing systems and pleremic writing systems, the third paper by Terry Joyce seeks to argue that morphographic, referring to orthographic units that primarily represent morphemes, is a preferable term than the conventional label of logographic, referring to orthographic units that primarily represent words. While touching on some of the implications of this terminological revision for how we think about writing and writing systems, Joyce's paper is mainly focused on the classification labels that have been applied to kanji as an element of the Japanese writing system. Joyce's paper concludes with a brief outline of some priming experiments for two-kanji compound words that have yielded findings that are consistent with the notion that morphological relationships reflected are within the mental lexicons of literate Japanese language users.

Imprecise terminology is also a major concern for the fourth paper by David Roberts. However, in the case of Roberts' paper, the terminology issues addressed are those that have hindered discussions of tone orthography and that have motivated Roberts to propose an explicit typology for tone orthographies that consists of six parameters. Defined in terms of choices, the six parameters are domain, target, symbol, position, density, and depth. While acknowledging the interdependent nature of these choices, Roberts also meticulously explains how the individual parameters can facilitate the orthographer in focusing on a particular aspect of an orthography for a tone language without losing sight of how the parameters interact to form a complex matrix of options. It is noteworthy that much of Roberts' discussion is devoted to the parameter of depth. Roberts' description of depth as being the most challenging parameter for a tone orthography typology would certainly seem to reflect the broader significance of the concept for all typologies of writing systems, as touched on briefly in the preceding review section and as discussed in Weingarten's paper.

The concept of orthographic depth is also central to the fifth paper by Martin Neef and Miriam Balestra. Drawing on the recoding model of graphematics proposed by Neef (2005), Neef and Balestra argue for a distinction between two interpretations of orthographic depth; graphematic transparency, related to the reliability of deriving the correct pronunciation of a word from its spelling, and orthographic transparency, related to the level of ambiguity associated with the conventional spelling of a word from its grammatical properties. More specifically, their paper focuses on graphematic transparency and outlines an approach to measuring it that yields a graphematic transparency value (gt-value). After a detailed explanation of calculating the gt-value for the German writing system, Neef and Balestra investigate the potential of this value as a meaningful comparison of different writing systems by describing the calculation of the gt-value for the Italian writing system. While the authors acknowledge the need to calculate gt-values for other writing systems in order to more fully understand the distribution of gt-values across different writing systems, their findings of a lower gt-value for Italian compared to the value for German is consistent with their graphematic analyses of German and Italian.

The final paper of this special issue is by Cláudia Silva, who presents an analysis of a Portuguese corpus of online chat conversations in terms of the observed deviations from conventional spellings. Silva argues that there are discernable patterns in the modifications as the online chat participants strive for more efficient ways of transcribing their language in writing, such that they are recreating certain orthographic features found in consonantal, syllabic, and morphographic writing systems. Silva's data would appear to further underscore Gelb's (1952)

claim that there are no 'pure' writing systems, which naturally raises serious issues about the nature of and the inherent limitations on the enterprise of seeking to develop meaningful typologies of writing systems. Moreover, given the inevitability of technological impacts on writing (historically including tool-related changes to letter shapes and the influence of the printing press on spelling conventions), the phenomenon of orthographic variation due to modern internet and communication technologies clearly warrants further investigation concerning the implications for the future evolution of writing systems.

Having briefly commented on the interesting contributions of each paper, we conclude this introduction to the special issue by expressing our deep appreciation to all the authors for their efforts and also to Martin Neef, general editor of Written Language and Literacy, for all his generous support throughout the editing process for this special issue, and by noting one more observation from Coulmas' (1996b: 1386) review of writing system typology:

> "Various typologies of writing systems have been suggested in the past and further typologies will no doubt be developed. Typologies are a means to create order in a complex and disorderly field. They are useful because they highlight problems in the study of writing and of language."

As guest editors of this special issue, our sincere hope is that typologies of writing systems to come can greatly benefit from the discussions of a wide range of problems intricately related to the study of writing and language that are addressed within the various papers of the special issue.

Notes

1.   Although Diringer's (1962) treatment of alphabetic writing has been criticized (Hill 1967; Coulmas 1996b), Diringer did acknowledge that alphabetic writing is technically a subdivision of phonetic writing, claiming that "alphabetic writing has within the past three thousand years assumed such importance as to deserve a category of its own" (Diringer 1962: 24). However, given that typologies should strive to be consistent in their classification criteria if they are to have value, the criticisms are certainly valid.

2.   Sampson (1994: 119–120) points out this division was intended to be more conjectural in nature, merely speculating on "whether there might ever be a semasiographic system comparable in expressive power to a spoken language", rather than arguing for the existence of such a system.

3.   The term *abjad* is formed from the first letters of the Arabic script, the most widespread example of this kind. The term *abugida* is an Ethiopic word formed from the initial letters according to a traditional ordering.

References

Coulmas, Florian (1992). Writing systems. In William Bright (ed.), *International encyclopedia of linguistics (vol. 4)*, 253–257. New York/Oxford: Oxford University Press.

Coulmas, Florian (1996a). *The Blackwell encyclopedia of writing systems*. Oxford: Blackwell.

Coulmas, Florian (1996b). Typology of writing systems. In Hartmut Günther & Otto Ludwig (eds.), *Schrift und Schriftlichkeit [Writing and its use] (vol. 2)*, 1380–1387. Berlin: De Gruyter.

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

Daniels, Peter T. (1996). The study of writing systems. In Peter T. Daniels & William Bright (eds.), *The world's writing systems*, 3–17. New York: Oxford University Press.

Daniels, Peter T. (2001). Writing systems. In Mark Aronoff & Janie Rees-Miller (eds.), *The handbook of linguistics*, 43–80. Oxford: Blackwell.

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

DeFrancis, John (2002). The ideographic myth. In Mary S. Erbaugh (ed.), *Difficult characters: Interdisciplinary studies of Chinese and Japanese writing* (*Pathways to advanced skills 6.*), 1–20. Columbus: National East Asian Language Resource Center, Ohio State University.

DeFrancis, John & J. Marshall Unger (1994). Rejoiner to Geoffrey Sampson, 'Chinese script and the diversity of writing systems'. *Linguistics* 32: 549–554.

Diringer, David (1962). *Writing*. London: Thames and Hudson.

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*, 111–134. Amsterdam: Benjamins.

Gelb, Ignace J. (1952). *A study of writing: The foundations of grammatology*. Chicago: University of Chicago Press.

Haas, William (1976). Writing: The basic options. In William Haas (ed.), *Writing without letters*, 131–208. Manchester: Manchester University Press.

Haas, William (1983). Determining the level of a script. In Florian Coulmas & Konrad Ehlich (eds.), *Writing in focus*, 15–29. Berlin: Mouton.

Hill, Archibald A. (1967). The typology of writing systems. In William M. Austin (ed.), *Papers in linguistics in honor of Leon Dostert*, 92–99. The Hague: Mouton.

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

Neef, Martin (2005). *Die Graphematik des Deutschen* (*Linguistische Arbeiten* 500). Tübingen: Niemeyer.

Rogers, Henry (2005). *Writing systems: A linguistic approach* (*Blackwell textbooks in linguistics* 18). Malden, MA/Oxford, UK: Blackwell Publishing.

Sampson, Geoffrey (1985). *Writing systems: A linguistic introduction*. Stanford: Stanford University Press.

Sampson, Geoffrey (1994). Chinese script and the diversity of writing systems. *Linguistics* 32: 117–132.

Sproat, Richard (2000). *A computational theory of writing systems*. Cambridge: Cambridge University Press.

Taylor, Isaac (1883). *The alphabet: An account of the origin and development of letters, 2 vols.* London: Kegan, Paul, Trench.

Trigger, Bruce G. (2004). Writing systems: a case study in cultural evolution. In Stephen D. Houston (ed.), *The first writing: Script invention as history and process,* 39–68. Cambridge: Cambridge University Press.

Unger, J. Marshall & John DeFrancis (1995). Logographic and semasiographic writing systems: A critique of Sampson's classification. In Insup Taylor & David R. Olson (eds.), *Script and literacy: Reading and learning to read alphabets, syllabaries and characters*, 45–58. Dordrecht: Kluwer Academic Publishers.

Authors' addresses

Terry Joyce
Tama University
School of Global Studies
802 Engyo, Fujisawa, Kanagawa, 252-0805
Japan

terry@tama.ac.jp

Susanne R. Borgwaldt
TU Braunschweig
German Department
Bienroder Weg 80
38106 Braunschweig
Germany

S.Borgwaldt@tu-bs.de