

Geoffrey Sampson

# Typology and the study of writing systems

DOI 10.1515/lingty-2016-0027

Received January 1, 2016; revised September 21, 2016

For the study of writing systems (still very much a minority branch of linguistics) the whole concept of typology is controversial, because there are influential scholars who believe that all the world's scripts are essentially of the same type. There has been a long though not particularly honourable tradition of discussing all writing systems as if they were more successful or less successful attempts to approximate the Roman alphabet, seen as the only possible ideal from which any other kind of script could only be viewed as a falling-off. In 1960, for instance, the distinguished anthropologist Sir Jack Goody and the literary critic Ian Watt co-authored a widely-read paper (Goody & Watt 1963) which used the term “literate societies” explicitly to mean societies using an alphabetic script, as opposed to societies like China which for over three millennia has (as Goody and Watt saw it) been struggling with a system of writing too crude to confer the benefits of literacy on the society which uses it. Michael & Jennifer Cole (2006: 305) note that Goody and Watt's paper, and subsequent related writings of Goody's, “have had an especially influential and continuing impact on a wide range of different disciplines [...] Goody's work on this topic continues to be used by anthropologists and historians, psychologists and sociologists”.

Scholars within the tradition described have often been prepared to recognize a difference in type between alphabetic scripts (the elements of which stand for segmental phonemes) and syllabic scripts (which divide speech into whole syllables). But this is a relatively minor distinction, set against the contrast between logographic scripts, which assign distinct marks to meaningful units of a language, i.e., words or morphemes, and phonographic scripts which represent phonological units of one size or another.<sup>1</sup> And even the syllabic v. alphabetic distinction was often blurred, for instance by lumping together, as “syllabic”,

---

<sup>1</sup> This paper does not aim to present a comprehensive typological classification of scripts. For that, see Sampson (2015: 20–39).

---

**Geoffrey Sampson**, Department of Linguistics and Modern Languages, University of South Africa, Theo van Wijk, 9-88, Muckleneuk Campus, Preller St, Pretoria, South Africa,  
E-mail: sampson@cantab.net

scripts in which the symbols for distinct syllables are graphically unrelated with scripts where series of syllables such as *ka ke ki* and *ba be bi* are written with constant outlines for /k/ or /b/ respectively, modified in consistent ways to show the particular following vowel. Furthermore, syllabic scripts tended to be described as if they were little more than temporary way-stations towards the ultimate ideal of alphabetic writing. The idea that a logographic script might be a fully-fledged, entirely satisfactory mode of written communication scarcely entered the purview of these scholars.

Up to a generation or two ago, such absurd views on the part of Western scholars could perhaps be explained as proceeding from widespread ignorance, even among professional linguists, about the details of how non-alphabetic scripts really work. But in 1989 John DeFrancis published a book subtitled “The diverse oneness of writing systems”, the central theme of which was that all scripts are indeed of a single, phonographic type. A few years earlier he had written that Chinese script “should be considered to be basically a phonetic system” (DeFrancis 1984: 125). DeFrancis certainly could not be dismissed as writing out of ignorance: the Chinese language was his special subject.

I have no hesitation, nevertheless, in saying that DeFrancis was quite wrong to suggest that all scripts are phonographic. I have refuted DeFrancis at length elsewhere (Sampson 1994, 2015: 21, 184–185), and I shall not repeat my detailed arguments here. The essence of DeFrancis’s mistake lay in failing to distinguish synchronic from diachronic modes of language description. It is reasonable to suggest that all scripts used as the normal written communication medium of a society were initially created as at least partly phonographic systems – I believe that is probably true (though it certainly is not true that early scripts were wholly phonographic). But it does not follow that all present-day scripts are phonographic, because a script, and the spoken language(s) it is used to represent, both change over time, and the result may be that the phonographic relationship is eventually lost. In the case of Chinese script that is exactly what happened. Chinese script used to write the modern Chinese language can only reasonably be described as a basically logographic script, even though from its long history the script does inherit features which often give limited and unreliable hints about present-day pronunciations.

DeFrancis makes much of the features of Chinese script just mentioned, but it is easy to demonstrate that a writing system need not be phonetically-based even to that limited extent. Chinese is not the only spoken language which is written using Chinese script. Japanese is written in a complex script, all elements of which ultimately derive from Chinese writing, and in particular the large share of the Japanese vocabulary which is native rather than borrowed and consists of lexical rather than grammatical morphemes is written with Chinese graphs for

translation-equivalents or near-equivalents. Since the two languages are genetically unrelated, there is no relationship whatever between the pronunciation of a native Japanese root and that of its Chinese translation-equivalent. Consequently, even when a Chinese graph does offer a good clue about its pronunciation IN CHINESE, it tells us nothing at all about its Japanese pronunciation.

For instance, the Chinese word for ‘taste’, *wèi*, is written by adding the graph for ‘mouth’, 口, to that for ‘not yet’, 未, giving 味; and in this particular case the ‘not yet’ element gives an excellent clue to the pronunciation of the compound graph, because in modern Mandarin ‘not yet’ is also *wèi*, a perfect homonym of ‘taste’. (This is an instance where the phonetic element may have become more rather than less appropriate over the millennia since the graphs were coined: Schuessler (2007: 512) shows the two words as having been only near-homonyms with slightly different vowels in the Middle Chinese period.) But, in Japanese, ‘taste’ is *aji* while ‘not yet’ is *mada* – the words are written with the same Chinese graphs, but their pronunciations are unrelated, so there is no phonetic basis at all for the structure of the ‘taste’ graph. That graph is purely a logogram, and this is the normal situation with respect to the writing of native Japanese vocabulary. Scripts really can be of different types (though, just as in the case of spoken-language typology, scripts commonly do not perfectly exemplify an ideal type).

It is ironic that there has been reluctance to recognize major typological differences between scripts, because, in reality, those differences seem to have more human significance than do typological differences among spoken languages. The contrast between an extreme case of inflecting languages, such as Greek, and an extreme isolating language, such as Vietnamese, is large in terms of the technicalities of formal language structure, but it is not usually thought to have large consequences for the functioning of the respective languages as vehicles of communication. Rightly or wrongly, the consensus appears to be that languages of different structural types carry out more or less the same tasks with similar efficiency though in somewhat different ways.

In the writing-systems domain, on the other hand, it is widely believed that differences between script types really matter to their users. For instance, there has been research on both Chinese and Japanese (summarized in Taylor & Taylor 1983: 404) suggesting that developmental dyslexia is strikingly less common for users of logographic scripts than for users of alphabetic scripts. More recently, McBride et al. (2015) do not discuss an absolute difference in frequency of incidence, but they argue that the syndromes covered by the general term “developmental dyslexia” are different in detail for users of the two types of script.

Ignatius Mattingly observed (1972: 144) that logographic scripts require a longer period to master than alphabetic scripts (because the number of distinct symbols to be learned is thousands rather than a few dozen), but mastering an alphabetic script is more intellectually challenging, because it requires the learner mentally to split up the physically-continuous speech stream into phonemes. William Hannas has argued eloquently that the former of these two points implies that logographic scripts impose a serious burden on societies which use them: “Instead of using language to learn, East Asians are wasting their youth and resources learning about language” (Hannas 1997: 125); “alphabetic literacy promotes creativity” (Hannas 2003: 5) whereas logographic script, Hannas believes, tends to stifle creativity, and in consequence he has argued that Chinese hopes of achieving First-World levels of economic development are doomed.

Hannas’s predictions look pretty foolish in view of the huge economic strides China has been making just in the few years since he was writing, and it would not be hard to construct an argument pointing in precisely the reverse direction. The Organization for Economic Co-operation and Development’s “PISA” programme is now giving us a triennial comparison of fifteen-year-olds’ educational attainments in basic subjects across 65 developed nations, and it is striking that countries with logographic scripts have been doing particularly well. The four highest scores both for reading and for science in the latest round of tests were for China, Hong Kong, Singapore, and Japan, and these territories were also among the top seven in maths.<sup>2</sup> These four territories use logographic script, whereas almost all the remaining 61 countries use phonographic script.

Measuring educational attainments across countries with different schooling systems, different spoken languages, and different cultural backgrounds involves huge problems of comparability, so I cannot claim that the implications of the PISA figures are as clearcut as they appear *prima facie*. And for that matter, I am sure that the last word has not been said about dyslexia among users of different script types, or about other issues in this area. But it is at least clear that there is room for serious, informed debate about various significant implications of script type for users’ lives and welfare.<sup>3</sup>

---

<sup>2</sup> The PISA programme uses Shanghai as a proxy for the People’s Republic of China as a whole.

<sup>3</sup> I argue elsewhere (Sampson forthcoming) that the history of scripts shows a trend away from simple phonemic systems towards more Chinese-like systems, and that this trend is an efficient response to social pressures.

So far as I know, the field of spoken-language typology has nothing like this. For instance, it is a cliché (in Britain, at least) that the societies of northern Europe are more economically dynamic than those of southern Europe, but I have never heard the least hint, either by economists or by linguists, that this might be partly explained by the fact that southern European languages are on the whole richer in inflections than the languages of northern Europe. The two sets of facts are taken (correctly, I would imagine) to be unrelated to one another.

It is true that linguistics has moved away from its earlier assumption that spoken-language typology is independent of the nature of the society using the respective language. In particular, Peter Trudgill (e.g., 2011) has shown that there are correlations between complexity of spoken-language structure and the “connectedness” of a society. But that does not imply that some types of spoken language do a better or worse job for their speakers, in the way that some types of script seem likely to be doing better or worse jobs for their users.<sup>4</sup>

So the study of writing systems is certainly not a branch of linguistics where typology has little significance. That said, there is a large problem about typological theorizing in this area. It is the same problem which creates difficulties for other scientific approaches to the study of writing systems: namely, there are few independent examples. The enterprise of science is about identifying general truths which apply across the board, and separating them out from the mass of properties which happen by chance to apply to particular instances of the phenomenon studied, but do not “have” to apply. In the domain of spoken languages, this is in principle straightforward to achieve. There are thousands of distinct spoken languages, belonging to many different families which appear to be genetically unrelated. (That is not to endorse the claims made by some linguists that there exist a rich range of “universals of language” – most of those claims have little basis in reality, as, e.g., Evans & Levinson (2009) have argued. But universal claims about spoken language are at least scientifically testable.) With writing systems it is different. All fully-alphabetic scripts (that is, with letters for vowels as well as for consonants) descend with only minor changes from the adaptation of some version of the Semitic alphabet, probably by a single individual Greek-speaker on a particular occasion, to write Greek. The only logographic scripts used to any serious extent in the modern world are Chinese script and its adaptations to write the languages of countries neighbouring China. It is as if we aimed to establish general theories about spoken

---

4 I say nothing here about correlations between spoken-language types and the script types used to write the respective languages, an issue I have discussed at length in Sampson (2015).

language in a world where the only languages spoken were the Romance languages plus Finnish and Hungarian. The result is that it becomes difficult to distinguish between facts which stem from the nature of human reading and writing behaviour, and facts which are mere matters of historical accident.

To take a simple hypothetical example: most alphabetic scripts use a single letter for the phoneme sequence /ks/ (Greek Ξ, Roman X). It is tempting to suppose that something about the phonetics of /ks/ makes it “natural” for that sequence to be treated as a single unit in writing. In reality, there is nothing natural about it; the existence of the letters Ξ and X is probably a pure historical accident. The Greek who first learned the alphabet must have struggled to interpret the alien sounds of a Semitic language in terms of the phonology of his own language, and perhaps came up with the interpretation /ks/ for a single Semitic sound.<sup>5</sup> Ever since then, speakers of Greek and of most languages written with the Roman alphabet have used a single letter for that pair of phonemes.

In this case, any serious hypothesis that /ks/ is written with a single letter because it is a natural phonological unit could be refuted by the fact that not all alphabetic orthographies treat it as such. The Cyrillic alphabet has no X equivalent, and some versions of the Roman alphabet do not use X even in “international” words (the Welsh for ‘taxi’ is *tacsi*). But, with so few unrelated scripts extant, there is no guarantee that in other, comparable cases we could find orthographies which make it clear that some apparent generalization about writing systems was spurious. A generalization which turned out to hold for each one of the thousands of spoken languages, on the other hand, could hardly be a mere coincidence. Hence it is easier in the case of spoken languages than in the case of scripts to establish that some group of properties genuinely belong together and define a natural type.

To summarize: relative to other branches of linguistics, for the study of writing systems issues of typology are unusually contentious, unusually significant, and also unusually difficult to research. That combination is perhaps unfortunate. But it is the way things are.

## References

- Cole, Michael & Jennifer Cole. 2006. Rethinking the Goody myth. In David R. Olson & Michael Cole (eds.), *Technology, literacy, and the evolution of society: Implications of the work of Jack Goody*, 305–324. Mahwah, NJ: Erlbaum.

---

<sup>5</sup> On the specifics of the relationship between sibilant letter-shapes and phonemes in Semitic languages and in Greek see Sampson (2015: 109).

- DeFrancis, John. 1984. *The Chinese language: Fact and fantasy*. Honolulu: University of Hawai'i Press.
- DeFrancis, John. 1989. *Visible speech: The diverse oneness of writing systems*. Honolulu: University of Hawai'i Press.
- Evans, Nicholas & Stephen C. Levinson. 2009. The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences* 32. 429–448.
- Goody, Jack & Ian Watt. 1963. The consequences of literacy. *Comparative Studies in Society and History* 5. 304–345. Reprinted in Jack Goody (ed.), *Literacy in traditional societies*, Cambridge: Cambridge University Press, 1968, and in Pier Paolo Giglioli (ed.), *Language and social context: Selected readings*, Harmondsworth: Penguin, 1972.
- Hannas, William C. 1997. *Asia's orthographic dilemma*. Honolulu: University of Hawai'i Press.
- Hannas, William C. 2003. *The writing on the wall: How Asian orthography curbs creativity*. Philadelphia: University of Pittsburgh Press.
- Mattingly, Ignatius G. 1972. Reading, the linguistic process, and linguistic awareness. In James F. Kavanagh & Ignatius G. Mattingly (eds.), *Language by ear and eye*, 133–147. Cambridge, MA: MIT Press.
- McBride, Catherine, Xiuhong Tong & Jianhong Mo. 2015. Developmental dyslexia in Chinese. In William S.-Y. Wang & Chaofen Sun (eds.), *The Oxford handbook of Chinese linguistics*, 688–696. Oxford: Oxford University Press.
- Sampson, Geoffrey. 1994. Chinese script and the diversity of writing systems. *Linguistics* 32. 117–132.
- Sampson, Geoffrey. 2015. *Writing systems*. 2nd edn. Sheffield: Equinox.
- Sampson, Geoffrey (forthcoming). From phonemic spelling to distinctive spelling. *Written Language and Literacy*.
- Schuessler, Axel. 2007. *ABC etymological dictionary of Old Chinese*. Honolulu: University of Hawai'i Press.
- Taylor, Insup & M. Martin Taylor. 1983. *The psychology of reading*. New York: Academic Press.
- Trudgill, Peter. 2011. *Sociolinguistic typology: Social determinants of linguistic complexity*. Oxford: Oxford University Press.