Areal script form patterns with Chinese characteristics

James Myers
National Chung Cheng University
http://personal.ccu.edu.tw/~lngmyers/

To appear in *Written Language & Literacy*

This study was made possible through a grant from Taiwan's Ministry of Science and Technology (103-2410-H-194-119-MY3). Iwano Mariko helped with katakana and Minju Kim with hangul, while Tsung-Ying Chen, Daniel Harbour, Sven Osterkamp, two anonymous reviewers and the special issue editors provided all sorts of useful suggestions as well.

Abstract

It has often been claimed that writing systems have formal grammars structurally analogous to those of spoken and signed phonology. This paper demonstrates one consequence of this analogy for Chinese script and the writing systems that it has influenced: as with phonology, areal script patterns include the borrowing of formal regularities, not just of formal elements or interpretive functions. Whether particular formal Chinese script regularities were borrowed, modified, or ignored also turns out not to depend on functional typology (in morphemic/syllabic Tangut script, moraic Japanese katakana, and featural/phonemic/syllabic Korean hangul) but on the benefits of making the borrowing system visually distinct from Chinese, the relative productivity of the regularities within Chinese character grammar, and the level at which the borrowing takes place.

Keywords: Chinese characters, Tangut script, Japanese katakana, Korean hangul, writing system grammar, script outer form, areal patterns

1. Areal phonological patterns and areal script patterns

Sinoform writing systems look Chinese, even when they are functionally quite different. The visual traits of non-logographic Japanese katakana and Korean hangul are nontrivially like those of logographic Chinese script, as are those of the logographic but structurally unique script of Tangut, an extinct Tibeto-Burman language of what is now north-central China. Sinoform scripts have long been recognized as sharing areal features (Unseth 2005), yielding what one might call a schriftbund (by analogy with the sprachbunds of Jakobson 1931). However, they also show that writing systems can borrow what Gelb (1963) called the outer form of a script without necessarily also borrowing what he called its inner structure or inner characteristics.

Given the importance of the writing-speech relationship, most studies on the Sinoform schriftbund have naturally focused on how written constituents represent spoken language (Tranter 2001, Miyake 2017a, Handel 2019). When outer form is discussed at all, it is either still in terms of its representational function, such as the characteristically Sinoform arrangement of written constituents within 'ideal squares' representing spoken constituents like morphemes or syllables (Tranter 2002), or else only with regard to broad formal features like the traditionally vertical axis of text lines or even "simply the calligraphic style or 'font'" (Tranter 2001, p. 185) of the ink brush.

Yet the borrowing of script outer form deserves more theoretical attention, since it shares deep similarities with borrowing in areal sound patterns. To appreciate these similarities, note first that "sound pattern" is a misnomer: what defines phonology, in contrast to morphology and syntax, is not physical sound but a system of formal linguistic regularities that are not directly interpreted semantically. After all, sign languages also have phonology (Brentari 2011), and more generally, phonology is at least as much mental as physical (Berent 2013). Since phonology is a system, the psychological processes that give rise to areal sound patterns are nontrivial. It is not just individual elements, like phonemes, that are borrowed, but also productive regularities like prosody (e.g. the influence of Bora syllable structure on Resígaro; Aikhenvald 2001) and phonological rules, like vowel harmony in Niger-Congo languages (Dimmendaal 2001) and possibly syllable-final devoicing across Europe (Blevins 2017). Thomason and Kaufman (1988) even argue that only words are literally "borrowed"; contact-induced changes in phonology or morphosyntax instead involve "interference through language shift" (p. 37), as non-fluent speakers of a language retain parts of their native grammars. Since borrowing systems must still conform to natural constraints on phonological processing and historical development, Blevins (2017) emphasizes that we should only ascribe cross-language similarities to borrowing if we rule out explanations based on universals (general linguistic tendencies) or inheritance (descent).

That writing systems also have grammar-like elements and regularities is an old idea, as reflected in the long pedigree of the term 'grapheme' (see critical reviews in Kohrt 1986, Daniels 2001, and a defense in Meletis 2019). Formal grammatical analyses have been offered for Roman letters (Watt 1975, Primus 2004), alphabetic spelling (Venezky 1967, Evertz 2018), abugida script (McCawley 1994), and Chinese characters (Wang 1983, Myers 2019). It is true that, in comparison to speaking and signing, writing is more an artificial tool than an intrinsic facet of human nature (Aronoff 1994), but the artificiality of writing should not be exaggerated. Readers and writers develop far more systematic knowledge than they are taught, as shown by numerous studies on children and adults (e.g. Pacton et al. 2001, Seidenberg et al. 1994 for alphabet scripts; Tsai & Nunes 2003, Lee et al. 2006, Myers 2019 for Chinese); see Meletis (2018) and citations therein for further arguments that writing is a natural linguistic modality.

Because alphabetic systems are closely linked to pronunciation, grammatical analyses of them tend to incorporate script/pronunciation interactions within a larger modality-independent phonological system (Wiese 2004, Fuhrhop et al. 2011, Evertz 2018). This approach is less useful in grammatical analyses of logographic scripts, where there is considerable systematicity that lacks any correspondence in speech: individual Chinese characters may represent the monosyllabic morphemes of spoken Chinese, but the complex character-internal patterns in stroke and constituent form obey only a written logic.

The vast majority of Chinese characters are semantic-phonetic, that is, made up of a component indicating the semantic class and a component representing the pronunciation, as illustrated in (1a), with pronunciations given in Hanyu Pinyin (the relevant aspect of each component is underlined). Most of the remaining decomposable characters are semantic compounds like that in (1b).

(1) a. 媽 *mā* 'mother' = 女 *nǚ* 'female' + 馬 *mǎ* 'horse'

b. 明 *míng* 'bright' = 日 *rì* 'sun' + 月 *yuè* 'moon'

While experiments confirm that readers actively interpret such components for meaning or pronunciation (as reviewed in Myers 2019, chapter 5), there are a great many other regularities in character-internal structure without any link to speech at all. Most of

these relate to strokes, including their axes and shapes, the direction and order in which they are written, how they may be combined to make well-formed constituents, and how they vary in size or shape in different contexts (see section 2). Adapting the modality-neutral definitions necessary for sign languages, Myers (2019, chapter 1) proposes the term 'character morphology' for regularities relating to constituent interpretation and 'character phonology' for those relating to pure form, noting that each share other properties traditionally associated with their namesakes in spoken and signed languages, for example that semantic components in semantic-phonetic characters act something like morphological affixes, and (following Wang 1983) that stroke alternations can be analyzed with something like phonological features and rules. Stroke-level lexical contrasts and regularities that cannot be reduced to general perceptual and motoric processes demonstrate that character phonology is distinct from 'character phonetics.' While Myers (2019) makes no claim that such an analysis is possible for all writing systems, Watt (1975) concluded that in the Roman alphabet, letters and strokes are analogous to morphemes and phonemes, respectively, because letters are composed of written sequences of strokes and only strokes can be decomposed into features; to this we can add that letters are also morpheme-like in being interpretable, albeit in pronunciation rather than meaning.

These considerations suggest that we can say something more nuanced about the shared "look" of Sinoform scripts than Tranter's (2001, p. 199) speculation that it is "due more than anything to the media used and the aesthetic tradition of the region." If Chinese characters truly have a script phonology, borrowing would involve formal regularities (rules and constraints) like those borrowed in areal sound patterns. Unseth (2005, p. 33) is close to drawing a similar conclusion when he compares cross-script borrowing to speakers using a "second language imperfectly, still maintaining habits from their first language"; in the case of early writers of Tangut script, katakana, and hangul, the "first language" was the Chinese script in which they were already fluent. Rather than mere "habits," however, I argue that what was borrowed was of much greater cognitive richness, namely a grammar.

To explore these ideas, this paper starts by reviewing the evidence for a number of patterns in Chinese character phonology (Section 2). It then looks for these patterns in a small but typologically varied sample of Sinoform scripts: Tangut script (Section 3.1), Japanese katakana (Section 3.2), and Korean hangul (Section 3.3). The conclusions in Section 4 attempt to explain why particular regularities were borrowed, modified, or ignored.

2. Chinese character phonology

Gelb (1963, p. 85) is not far wrong about Chinese script when he writes that "in its outer form the writing has changed greatly in the course of its long history, but from the point of inner characteristics the oldest inscriptions hardly differ from those of recent times." Admittedly, there have also been changes in character morphology (as defined above) from their origins in oracle bones before 1200 BCE up to modern standard script, mature by 600 CE (Qiu 2000), as seen today in the traditional system retained in Hong Kong, Macau, and Taiwan and in the simplified system developed in the People's Republic of China (Wiedenhof 2017). However, these morphological changes have been far less dramatic than changes in character phonology, most obviously the replacement of the curves of older systems with today's straight strokes (Myers 2019, section 4.3.1.2). Such changes in outer form cannot be explained solely by the writing media as Tranter (2001) implies, given that ink brushes are also used to produce nonlinear calligraphic Chinese styles and the circles of Korean hangul. Nor is it helpful to ascribe character outer form to some amorphous aesthetic tradition, since like a phonological system, it has a productive internal logic of its own (Wang 1983, Myers 2019). Indeed, while the patterns discussed in this section will be illustrated

with a single traditional Chinese character typeface, chosen because it most closely mimics careful (non-cursive) brush pen handwriting, they also appear in simplified characters, most other modern typefaces, and ordinary handwriting (see Myers 2019, sections 4.3.2.2, 4.4.1, and 4.4.2, respectively). It is in fact a striking sign of the abstractness of Chinese character phonology that purely formal patterns with originally perceptual and motoric motivations (see below) remain obligatory even in mechanical printing.

To start with, the stroke inventory, as exemplified in Figure 1, is not arbitrary but derivable from a few simple principles (see also Wang 1983, Peng 2017): strokes are straight by default, fall primarily along the two cardinal axes (horizontal and vertical) and otherwise along the two major oblique axes (diagonals), with each stroke written from left to right and/or from top to bottom. These principles correctly predict that for three of the axes (|, –, \) there is only one stroke type, but for the fourth (/) there are two: rightward/upward and leftward/bottomward. This featural system also predicts a stroke without a lexically specified axis, the dot, by default drawn as a very short falling diagonal (\) to conform with both directional principles. These basic strokes may then be linked (endpoint to starting point) to produce a variety of complex strokes.

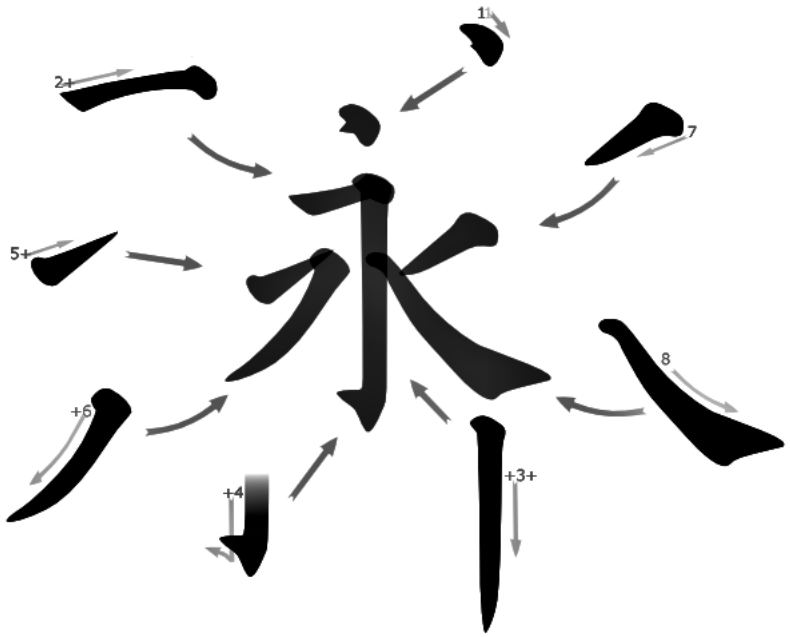

Figure 1. Axes, directions (arrows), order (digits), and features (e.g. hooking) of basic strokes (e.g. <一>) and complex strokes (e.g. <㇆>) in <永> *yǒng* ‘forever’[1]

The principles themselves seem to be motivated by perception and motor control: writing systems universally favor the two cardinal axes (Changizi et al. 2006) and the rightward and bottomward directions allows a right-handed writer to pull rather than push the writing instrument (Watt 2015), a preference also shown by non-Chinese, right-handed children asked to copy simple geometric figures (Ninio and Lieblich 1976).

Variations in dot axis also have clear external motivations, particularly visual symmetry (another orthographic universal; Morin 2018). As illustrated in (2), dots at opposite ends of rows have mirror-image axes.

[1] Derived from public domain image at https://commons.wikimedia.org/wiki/File:8_strokes_of_%E6%B0%B8-zh.png.

(2) 首 *shǒu* ‘head’ 平 *píng* ‘level’ 當 *dāng* ‘to face’
受 *shòu* ‘accept’ 只 *zhǐ* ‘only’ 黑 *hēi* ‘black’

Universal principles help explain how strokes are combined to form Chinese character constituents. Cardinal and oblique axes were observed to avoid mixing in the 116 non-logographic scripts analyzed by Morin (2018), and the same is true of Chinese constituent-internal strokes, which are frequently parallel, as in (3a), or orthogonal, as in (3b). Moreover, when one stroke touches another without crossing it, it is generally at the initial point of the second-written stroke. Combined with the directional principles, means that stroke combinations like ┬ and ├, as in (3c), are more frequent than those like ┴ and ┤, another preference shared with the children tested by Ninio and Lieblich (1976).

(3) a. 三 *sān* ‘three’ 川 *chuān* ‘river’ 彡 *shān* ‘hair’
b. 十 *shí* ‘ten’ 乂 *yì* ‘govern’
c. 下 *xià* ‘below’ 人 *rén* ‘person’

Character constituents also show several more distinctively Chinese stroke regularities. The lowest or rightmost element within a constituent is almost always enlarged, whether the affected component is a single stroke as in (4a) or a stroke group as in (4b). Lexical exceptions, like those in (4c), are rare, often to mark lexical contrast (compare with first two characters in (4a)); as Myers (2019, section 3.5) points out, character phonology shows the same diagnostics of lexical phonology as in spoken and signed languages (Hargus & Kaisse 1993). Since enlarged elements appear at the bottom or right, they are written last, suggesting that the external motivation is the same motor-control force that lengthens final syllables in both spoken and signed languages (see, respectively, Beckman & Edwards 1990 and Sandler 1993). In other scripts this effect is just barely visible, as in the optional and barely discernible enlargement of the lower portion of <B>, but in Chinese it has been ‘phonologized’. The enlargement of entire Chinese constituents also argues against its remaining a mere motoric effect, and even when it affects individual strokes, enlargement occurs in positions defined not by the ‘ideal square’, but by internal constituent boundaries, as illustrated in (4d); that is, it is sensitive to character morphology. Some characters seem to show secondary enlargement at the top or left, as in (4e), but this is not consistent, as shown in (4f); secondary leftmost enlargement may relate to curving and secondary topmost enlargement to the similarly inconsistent enlargement at the top in triple reduplication, as in (4g). In any case, Myers (2019, section 5.3.3) provides experimental evidence that Chinese readers judge nonce stroke combinations as more character-like if the longest of a set of parallel strokes appears at the bottom or right.

(4) a. 土 *tǔ* ‘earth’ 未 *wèi* ‘not yet’ 手 *shǒu* ‘hand’
川 *chuān* ‘river’ 州 *zhōu* ‘state’
b. 炎 *yán* ‘inflame’ 串 *chuàn* ‘string’ 飛 *fēi* ‘fly’
林 *lín* ‘forest’ 比 *bǐ* ‘compare’
c. 士 *shì* ‘gentleman’ 末 *mò* ‘final’
d. 圭 *guī* ‘jade tablet’ (cf. 土 in (4a))
e. 三 *sān* ‘three’ 王 *wáng* ‘king’ 川 *chuān* ‘river’
f. 彡 *shān* ‘hair’ 拜 *bài* ‘to bow’ 爪 *zhuǎ* ‘claw’

g. 森 *sēn* ‘profusion’ (cf. sans serif 森) 垚 (variant of 澀 sè ‘tart’)

Another distinctively Chinese constituent-level pattern is the curving of the leftmost vertical stroke, already seen in some of the above examples. Further examples are given in (5a); the last two show that the rule overrides mirror reversal. Vertical strokes are never curved when not at the left edge, as illustrated by the straight strokes in the center and at the right in (5b). However, curving has many more exceptions than enlargement, as in (5c). As first observed by Wang (1983) and confirmed quantitatively by Myers (2019, section 3.4.5), curving is more likely in tall thin constituents (like <月> in (5a)) than in wide ones (like <冊> in (5c)), but this remains a mere tendency. The greater degree of lexicalization in curving is also suggested by examples like those in (5d), which anomalously show curving at the left edge of the whole character rather than at the left edge of an individual constituent. Unsurprisingly, then, curving showed weaker generalizability to nonce stroke groups than enlargement in the character-likeness experiments reported in Myers (2019, section 5.3.3).

(5) a. 介 *jiè* ‘introduce’ 非 *fēi* ‘not’ 月 *yuè* ‘moon’ 周 *zhōu* ‘to circle’
片 *piàn* ‘slice’ 爿 *pán* ‘slice’ (dialectal classifier)

b. 中 *zhōng* ‘middle’ 平 *píng* ‘level’ 年 *nián* ‘year’
耳 *ěr* ‘ear’ 斗 *dòu* ‘dipper’

c. 門 *mén* ‘door’ 冊 *cè* ‘book volume’ 同 *tóng* ‘same’

d. 羚 *líng* ‘antelope’ (= 羊 *yáng* ‘sheep’ + 令 *lìng* ‘command’)

Still less productive is hooking at the end of a stroke. Hooking into the concavity of a deeply bent stroke is virtually obligatory, as in (6a), but its very obligatoriness allows it to be analyzed as an intrinsic feature of this stroke type rather than something added by productive rule. As illustrated in (6b), rightward hooking on vertical strokes appears in another suspiciously specific environment, namely to the immediate left of crossed strokes; Myers (2019, section 4.3.1.2) provides historical evidence that this pattern may have spread by analogy from the first two examples. Hooks on the end of horizontal strokes are plausibly analyzed as short strokes, since uniquely they may form a chain with another stroke, as in <了> in (6c), or even alternate with full strokes, as in the final example in (6c). Leftward hooking on vertical strokes is nearly lexically contrastive, as seen by the near-minimal pairs in (6d), though it never appears on the left edge of a constituent and, as first observed by Wang (1983) and confirmed quantitatively by Myers (2019, section 3.4.6), is more common when the stroke is topped with other material, as in the first set in (6e). These leftward hooking generalizations are quite weak, however, and again the character-likeness experiments in Myers (2019, section 5.3.3) failed to provide unambiguous support for their productivity.

(6) a. 豸 *zhì* ‘wormlike invertebrate’ 司 *sī* ‘control’ 心 *xīn* ‘heart’

b. 氏 *shì* ‘clan’ 民 *mín* ‘folk’
長 *cháng* ‘long’ 良 *liáng* ‘good’ 衣 *yī* ‘clothing’

c. 疋 *shū* ‘foot’ 也 *yě* ‘also’ 了 *le* ‘completive aspect’
當 *dāng* ‘to face’ (= 尚 *shàng* ‘yet’ + 田 *tián* [facing] ‘field’)

d. 事 *shì* ‘thing’ 聿 *yù* ‘writing brush’ (archaic)
乎 *hū* ‘at’ 平 *píng* ‘level’
e. 可 *kě* ‘can’ 牙 *yá* ‘tooth’ 手 *shǒu* ‘hand’
小 *xiǎo* ‘small’ 才 *cái* ‘ability’ 水 *shuǐ* ‘water’

By contrast, stroke patterns that apply across constituents are highly productive. Regardless of constituent function as semantic or phonetic component, the lowermost horizontal stroke is diagonalized when followed by another constituent to its right, as in (7a), even when part of a complex stroke, as in (7b). The lower right falling diagonal becomes a dot in the same environment, as in (7c). Both alternations help shorten the writing distance to the following constituent (from the lower right of the first to the upper left of the next) but are not synchronically reducible to this motoric process (i.e., they are part of character phonology rather than character phonetics), as demonstrated by the reversal in stroke order in (7d): the vertical stroke is written last when the constituent is on its own, but when bound to another it is the diagonal stroke that is written last.

(7) a. 土 *tǔ* ‘earth’ 地 *dì* ‘ground’
b. 九 *jiǔ* ‘nine’ 鳩 *jiū* ‘pigeon’
c. 夫 *fū* ‘husband’ 規 *guī* ‘rule’
d. 牛 *niú* ‘ox’ 物 *wù* ‘thing’

3. Borrowing, modifying, and ignoring formal Chinese character regularities

Having identified a variety of formal patterns in Chinese characters (principles of strokes and their combinations, dot axis, enlargement, curving, hooking, diagonalization, and dotting), we now examine whether they have been borrowed into three otherwise quite different Sinoform scripts, all originating well after Chinese characters had developed their modern standard form: Tangut script, Japanese katakana, and Korean hangul.

3.1 Tangut script

Unlike standard Chinese script and Japanese katakana, but like several other Sinoform scripts (Song 1998), the Tangut script did not develop naturally over centuries, but was instead imposed by government fiat. It was first promoted in 1036 CE by the ruler of the Tangut Empire (just west of the Song Empire), remained in limited use even after the Tanguts were conquered by the Mongols in 1227, but was then almost completely forgotten until its rediscovery in the nineteenth century; even today analytical questions remain (Gong 1982, Kychanov 1996, Galambos 2015).

Roots in spoken Tangut tended to be monosyllabic (Miyake 2017b), making a Chinese-inspired logographic script a good fit. At the same time, the social and political desire to distinguish themselves from the Song led the Tanguts to make their script differ much more from standard Chinese than more familiar Sinoform logographic systems like Japanese kanji. The script morphology of Tangut characters, though sometimes borrowed directly from Chinese (Gong 1982), was more likely to involve semantic compounding than the affixation of semantic elements to phonetic constituents (the dominant operation in Chinese character morphology), and Tangut constituents were themselves almost always morphologically complex, forcing the deletion of one or more components when compounded with others (Tranter 2001). Given such complexities, Miyake (2017a: 37)

expresses the consensus view (see also Kychanov 1996) that "[t]he true structure of many Tangut characters is obscure."

The script phonology of Tangut is much more straightforward.[2] In the standard script (which was used both in woodblock printing and in formal handwriting, as opposed to cursive), stroke groups are composed of mostly straight lines along the cardinal and major oblique axes, including complex strokes linking basic strokes into chains, with the occasional curve or hook. As suggested in (8a), the oblique axis and complex strokes are used more frequently than in standard Chinese script (both are common in cursive Chinese styles, but the Tangut examples here are not cursive). Cardinal and oblique axes still tend not to mix, and dot axis is also like that in Chinese in lying along the main diagonal by default, as seen in (8b), and obeying symmetry in rows of dots, as in (8c). Since all of these similarities conform directly to universal principles, they need not reflect borrowing.

(8) a.  *tạ* 'wrinkle'  *ɣạ* 'mad'  *ljwɨj* 'great great grandson'

b.  *rjijr* 'nearby'  *zjir* 'back'

c.  *khjow* 'respectability'  *bji* 'below'  *tsow* 'soup'

More distinctively Chinese is Tangut's enlargement of the lowest of a set of horizontal or diagonal strokes, as in (9a), and of the rightmost of a set of vertical strokes, as in (9b).

(9) a.  *zji* 'a place name'  *ɣjɨ* 'ore'  *rjɨj* 'pavilion'

b.  *nu* 'to violate'  *śio* 'to mate'

Some of the above examples also show curving in the leftmost of a set of vertical strokes; further examples are given in (10a). This left-edge position seems to be defined at the level of the constituent, not character, as shown in (10b). These aspects of curving in Tangut thus appear to have involved genuine borrowing from Chinese character phonology.

(10) a.  *tshjɨj* 'intimate'  *xə* 'to seek'

b.  *mẹ* 'to lie'  *pjạ* 'dregs'

Also as in Chinese script, the leftmost vertical stroke is not always curved, but in Tangut the exceptions are much more predictable: a stroke in this position is consistently straight when topped by a horizontal stroke, as in (11a), or when a horizontal stroke contacts or crosses it, as in (11b). These restrictions may be generalized from similar but less robust tendencies seen in Chinese characters like those in (12a) and (12b), respectively.

[2] The typeface used here is Tangut Yinchuan Version 12.000, expanded by Andrew West (https://www.babelstone.co.uk/Fonts/Yinchuan.html) from the XXZT typeface created for Li (2008) by Jing Yongshi, who holds the copyright. Its appearance matches that of texts in historical facsimiles (see https://www.babelstone.co.uk/Tangut/TongyinLookup.html, https://www.babelstone.co.uk/Tangut/WenhaiLookup.html). Pronunciations and English glosses are from Li (1997, 2008) as indexed by Andrew West at (https://www.babelstone.co.uk/Tangut/XHZD_Index.html, https://www.babelstone.co.uk/Tangut/XHZD_EnglishIndex.html).

(11) a. *xwa* ‘trench’ *tśhjɨ* ‘narrow’ *tjụ* ‘saliva’

b. *tji* ‘east’ *swa* ‘drop’ *wjụ* ‘dung’

(12) a. 同 *tóng* ‘same’ (cf. 周 *zhōu* ‘to circle’)

b. 非 *fēi* ‘not’ (cf. variant 非)

By contrast, hooking has not been productively borrowed as a process, appearing only within concave strokes, as in (13), with the hooked stroke in (13a) taken directly from the Chinese inventory, as in (14a), and that in the left-edge constituent in (13b) being a modification of the Chinese stroke in (14b).

(13) a. *do* ‘poison’

b. *śjwa* ‘to hold’

(14) a. 乚 (as in 化 *huà* ‘change’)

b. ㇆ (as in 刀 *dāo* ‘knife’)

Finally, despite the fact that Tangut characters are virtually always made up of more than one constituent, there is no sign of the cross-constituent stroke alternations of diagonalization and dotting. However, this may simply be because neither has a chance to apply: left-edge constituents never have horizontal strokes at the bottom, perhaps to keep them thinner (e.g., the top-edge constituents in (9a) above never appear at the left), and falling diagonal strokes at the lower right of a constituent are always part of a cross, as in (15a), where dotting is also blocked in Chinese script, as shown in (15b).

(15) a. *gjɨ̱* ‘nine’

b. 女 *nǚ* ‘female’ 姐 *jiě* ‘older sister’

3.2 Japanese katakana

By the end of the 700s CE, following a similar practice in Korea, Japanese writers were annotating Chinese text with extremely simplified forms of Chinese characters to indicate Japanese pronunciations (Whitman 2011). These evolved into two separate semi-syllabic (moraic) systems in modern Japanese script. The more common is hiragana, used for affixes, function morphemes, and native lexical items lacking kanji, while katakana is used primarily for transliteration in loans and onomatopoeia. This section focuses on katakana because it has a script phonology derived from that of standard Chinese script, unlike hiragana, which is derived from cursive Chinese calligraphy. Cursive is not merely a script-phonetic implementation of standard script but has its own grammar, with conventionalized rules that can make it illegible to standard script readers (Tseng 1993). A proper grammatical analysis of hiragana’s development would thus require us to first provide a grammar of Chinese cursive, which space precludes here.

The modern form of katakana developed naturally over a millennium (Tsukishima 1977, 1997). Hansell (2002: 166) claims that in its derivation from Chinese characters “the simplifications are usually the loss of one or a few strokes, no more radical than many of those involved in the simplification of Chinese that has occurred in Mainland China” and that

"there is no historical or structural principle" to distinguish them from kanji; thus "their only difference is the respective functions they perform within Japanese." This claim is misleading in two ways. First, as Song (1998: 23) notes, the changes were actually often so great that "the origins can not be judged with certainty by the outer forms alone." Second, in contrast to the nearly full preservation of Chinese character grammar in simplified characters (Myers 2019, section 4.3.2), modern katakana has retained only some of the formal patterns reviewed in section 2, while developing new regularities of its own.

The stroke inventory in katakana is quite similar to that in standard Chinese script, including the preference for cardinal axes and the tendency to avoid mixing them with oblique axes, but even at this most fundamental of formal levels, katakana still goes its own way. As noted earlier, Chinese script has two strokes that differ only in writing direction, but they never form lexical contrasts. However, katakana has not one but two pairs of signs differing only in these stroke types: the diagonal strokes in (16a) are written upward and rightward, while those in (16b) are written downward and leftward. The pronunciations (in standard Hepburn Romanization) and source characters (from Tsukishima 1977, 1997) confirm that neither explains the modern forms: the stroke direction contrast does not represent any pronunciation feature, and none of the possible source characters have upward/rightward strokes. The two strokes also differ slightly in axis, the only feature visible in sans serif fonts (illustrated in square brackets in (16)), but Chinese characters only make such physically subtle distinctions between the conceptually quite distinct oblique and cardinal axes, as in (17). Meanwhile, the complex stroke that forms (16c) is entirely absent in standard Chinese script, unsurprisingly, perhaps, given that it comes from cursive Chinese script, and indeed is the only katakana form identical to its hiragana counterpart.

(16) a. ン *-n* (< 尓?) [ン] シ *shi* (< 之) [シ]
b. ソ *so* (< 曾) [ソ] ツ *tsu* (< 津, 川, 州) [ツ]
c. ヘ *he* (< cursive form of 阝 in 部)

(17) a. 千 *qiān* 'thousand'
b. 干 *gān* 'to concern'

The two katakana in (18a) have strokes slightly off from true horizontal and vertical, perhaps due to the inheritance of the off-horizontal stroke from one source character and the off-vertical stroke from the other, combined with the universal principle of stroke orthogonality. Just as in Chinese script, however, these odd axes do not contrast directly with true horizontals and verticals, since the vertical strokes in (18b) are part of complex strokes.

(18) a. ヤ *ya* (< 也) キ *ki* (< 幾)
b. セ *se* (< 世?) モ *mo* (< 毛)

Some of the above examples show that the default falling diagonal dot axis is the same in katakana as in Chinese script, as does the voicing diacritic illustrated in (19a), which may reflect either borrowing or universal motoric motivations. By contrast, the diacritic in (19b), indicating that what otherwise would start with /h/ should here start with /p/, is un-Chinese-like in being circular, though the same mark has long been used in Chinese as the punctuation in (19c).

(19) a. タ *ta* ダ *da*
b. ハ *ha* パ *pa*
c. ◦ (period/full stop in modern usage)

The only instance of dot axis symmetry, in (20a), is inherited directly from the source character, though the reemergence of default dot direction in (20b) with the loss of its source's adjacent parallel stroke may hint at productivity.

(20) a. ホ *ho* (< 保)
b. ト *to* (< 止)

Enlargement of the lowermost and rightmost of a set of parallel strokes is also observed in katakana, but this is virtually always inherited directly from the source character, as shown in (21a-b). Even the apparently productive enlargement in (21c) may merely be due to ad hoc analogy with (21b), which emerged much earlier (based on the indicated years, from Tsukishima 1977, in which they first reached their modern forms).

(21) a. エ *e* (< 江) チ *chi* (< 千) テ *te* (< 天) ニ *ni* (< 二)
ミ *mi* (< 三) モ *mo* (< 毛)
b. リ *ri* (< 利) (810)
c. サ *sa* (< 散) (1126)

Vertical stroke curving seems quite regular in katakana, but it works in almost the opposite way from Chinese script: only once does it occur at the left edge, in (22a), inherited from its source character, and otherwise it occurs at the right edge, as in (22b), originally via reanalysis of a hook in its source character, or centrally, as in (22c), where the dates (from Tsukishima 1977) suggest that this uniquely katakana pattern was sparked in the first two forms (テ *te* and ナ *na*) via the inheritance of the leftmost curved stroke in the source characters, with the loss of other strokes shifting it to the center. However, these were only established over two hundred years after katakana's origins, and the process took another three hundred years to complete. This history suggests that the Chinese left-edge curving rule, which still needed to be applied in kanji, remained a strong hindrance to the development of the katakana central curving rule in the minds of Japanese writers.

(22) a. ル *ru* (< 流)
b. リ *ri* (< 利) サ *sa* (< 散)
c. テ *te* (< 天) (1020)
ナ *na* (< 奈) (1020)
チ *chi* (< 千) (1073)
ケ *ke* (< 介, 个) (1196)
ア *a* (< 阿) (1344)

The remaining Chinese character patterns show even less productivity in katakana. Hooking in convex strokes is absent, as seen in (23), in fact dropped from its source in (23a). Hooking on (off-)horizontal strokes is either inherited directly from the

source character, as in (24a), or by apparent analogy to this, as in (24b), with the very long delay (based on the dates in Tsukishima 1977) suggesting a lack of productivity. Moreover, consistent with the analysis of this hook type as actually being a short stroke in Chinese script, earlier versions of the katakana form in (24b) looked more like its hiragana counterpart in (24c), with a separate short stroke rather than an attached hook; this hook can also optionally make contact with the following stroke, as in (24d) ((24e) shows a minimal pair involving this stroke combination, though the second katakana is not used in the contemporary standard orthography). Hooks on vertical strokes were simply inherited from the source characters or constituents, as in (25). The cross-constituent diagonalization and dotting alternations have no chance to apply at all, since each katakana fills up its own ideal square.

(23) a. ヒ *hi* (< 比)
b. セ *se* (< 世?)

(24) a. ヤ *ya* (< 也) (883)
b. セ *se* (< 世?) (1697)
c. せ *se*
d. ア *a* (cf. sans serif variant 𠮟)
e. エ *e* 　ヱ *we*

(25) a. カ *ka* (< 加)
b. オ *o* (< [obsolete variant of 於])
c. ホ *ho* (< 保) (cf. 茶 *chá* ‘tea’ and variant 茶)

3.3 Korean hangul

Hangul was presented to his people by King Sejong the Great in 1446 CE. However, in stark contrast to Tangut script, also created by government fiat, hangul is often praised for its great elegance (e.g. Lee 2009). Simplifying somewhat, each ideal square represents a syllable, each stroke group within the square represents a phoneme (a letter), and the strokes themselves provide information about spoken phonological features (e.g. vowel versus consonant, or place and manner of articulation).

These internal characteristics have remained quite stable despite large changes in the outer form of hangul letters (script phonology) that began “within a decade of their promulgation” (King 1996: 225). A change of writing medium was involved, as the simple geometric forms of the woodblock-printed foundation documents (circles, round dots, and lines along the cardinal and major oblique axes) had to be realized in brush pen handwriting (Ahn 2017). This involved the adoption of the Chinese stroke directions and orders, though since these are universally motivated, as we saw earlier, this need not have involved direct borrowing.

As with the other systems discussed above, modern hangul does not simply preserve all of the formal regularities of standard Chinese script, but rather borrows some while modifying or ignoring others. The most obvious non-Chinese feature is the circularity in the letters in (26), shown along with their pronunciations in Revised Romanization of Korean.[3]

[3] IPA is included with some later examples where the romanization is insufficiently clear.

(26) a. ㅇ *∅-/-ng*
b. ㅎ *h*

Dots are no longer circular in modern hangul, but they do not take on the Chinese default falling diagonal axis either. Instead they are realized as short strokes orthogonal to and contacting the adjacent stroke, as illustrated in (27a); compare the short horizontal stroke in the first example with the falling diagonal dot in the Chinese character in (28). Depending on the typeface, the short stroke in the modern hangul in (27b) may also be realized as parallel to the adjacent stroke (though not necessarily in rapid handwriting). Note also that the geometric logic of hangul's design forces systematic violations of the universal disfavoring the ㅗ and ㅓ stroke configurations.

(27) a. ㅏ *a* ㅓ *eo* [ʌ] ㅗ *o* ㅜ *u*
b. ㅎ *h* (cf. variant ㅀ)

(28) ㅏ *bǔ* 'to divine'

As pointed out by Myers (2019, section 3.4.4), hangul has borrowed the enlargement regularity from Chinese script, as seen above and in (29). In fact, the lower of the two horizontal strokes in (29a) was already lengthened relative to the upper one even in its original geometrical form (see facsimile in Kim 2005). That what is borrowed is an abstract regularity rather than specific stroke types is suggested by the enlargement of the lower vertical strokes in (29b) even without a clear contrast in horizontal stroke size. The similarly shaped Chinese characters in (30a) cannot show enlargement contrasts due to the need for each to fill up an ideal square (which individual hangul need not do, as seen in (29b)). The closest Chinese script gets to the stroke length contrast in (29b) is the character pair in (30b), though even here it seems that bottommost enlargement may be targeting the box-shaped stroke group rather than the vertical stroke.

(29) a. ㅍ *p*
b. ㅗ *o* ㅜ *u* ㅛ *yo* ㅠ *yu*

(30) a. 上 *shàng* 'above' 下 *xià* 'below'
b. 由 *yóu* 'cause' 甲 *jiǎ* 'shell'

In hangul, however, enlargement is sensitive to stroke axis: only vertical strokes show enlargement on the right, as in (31a), whereas the short horizontal strokes (historically dots) are enlarged on the left instead, as in (31b). This latter pattern not only differs from Chinese script but also goes against the general motoric tendency towards final-gesture enhancement.

(31) a. ㅐ *ae* [ɛ] ㅒ *yae* [jɛ] ㅔ *e* ㅖ *ye*
b. ㅏ *a* ㅓ *eo* [ʌ] ㅑ *ya* ㅕ *yeo* [jʌ]

In contrast to the borrowing (with modification) of enlargement, there is no curving or hooking in hangul. This may relate to differences in the relative productivity of these three patterns in Chinese script. Enlargement may have been carried over without much thought, as

allophonic patterns are carried over in second language speech, whereas curving and hooking, being more lexically specified in the source script, may have been less mentally active when writing in the new script.

The combination of hangul letters into syllables within an ideal square provides ample opportunity for the cross-constituent alternations of diagonalization and dotting to apply, and indeed both do, just as in Chinese script (Myers 2019, section 3.3.1.2). While illustrated with simple letter combinations in (32), these generalizations also apply in full syllabic ideal squares. Diagonalization is seen in letter combinations representing diphthongs, as in (32a-c), and consonant clusters, as in (32d-e), where the affected horizontal element is part of a complex L-shaped stroke. Dotting is seen in the doubled letters representing coda consonants with so-called tense or fortis phonation, as in (32f-g), though in my example font it is somewhat clearer in (32f). Note that the application of diagonalization in (32b) also reconfirms that this pattern is not merely the mechanical effect of the efficient movement of a writing instrument, since due to the directional principles, the affected stroke is not the last one written within its letter.

(32) a. ㅗ *o* 괴 *oe* [we] 과 *wa* 괘 *wae* [wɛ]
b. ㅜ *u* 귀 *wi* 궤 *we* 궈 *wo*
c. ㅡ *eu* [ɯ] 긔 *ui* [ɯi]
d. ㄴ *n* 냐 *nj* 낭 *nh*
e. 르 *l /r* 랑 *lg* 램 *lm* 랩 *lb* 랏 *ls* 랕 *lt* 랖 *lp* 랗 *lh*
f. ㅅ *s* ㅆ *ss*
g. ㅈ *j* ㅉ *jj*

4. Discussion

It has long been noted that the four Sinoform systems discussed here (Chinese script, Tangut script, Japanese katakana, and Korean hangul) share formal properties, but observations have generally been restricted to the ideal square, traditionally vertical text lines, the stroke inventory, and the mechanical effects of ink-brush calligraphy. As reviewed in Table 1, however, more careful analysis reveals a number of deeper similarities and differences (illustrated with the indicated examples). Some formal regularities of Chinese script are inapplicable in the derivative scripts and some are ignored; some may be shared solely due to universal visual or motoric constraints and some are directly inherited from source characters or strokes as elements rather than as regularities; some are modified (or even systematically reversed); and some are unambiguously borrowed unchanged from Chinese script grammar.

Table 1. Chinese character phonology in some non-Chinese Sinoform scripts.

	Tangut	Katakana	Hangul
Left-to-right and top-to-bottom stroke direction and order	Unknown (but probably universal)	Universal	Universal
Cardinal axis preferred (3)	Universal (with modification: more oblique) (8)	Universal (with modification: off-vertical and off-horizontal axes) (cf. 16, 18a)	Universal (27a)
Avoid mixing cardinal and oblique axes (3)	Universal (with modification: more mixing) (8)	Universal	Universal (27a)
Default falling diagonal dot axis (Figure 1)	Borrowed or universal (8b)	Borrowed or universal (19)	Modified (dots instead orthogonal or parallel to adjacent stroke) (27)
Dot axis symmetry (2)	Universal (8c)	Inherited (from source character) (20a)	Not applicable
Stroke contact at start of stroke (3c)	Universal (8)	Universal	Universal
Enlargement at bottom (4a, b) vs. (4c)	Borrowed (9a)	Inherited (from source characters) (21a)	Borrowed (29)
Enlargement at right (4a, b)	Borrowed (9b)	Mostly inherited (from source characters) (21b, c)	Modified (only applies to vertical strokes; horizontal strokes enlarge on left) (31)
Curving of leftmost vertical stroke (5a) vs. (5c, 12)	Borrowed (with modified restrictions) (9b, 10) vs. (11)	Modified (reversed contexts: curving mostly everywhere except left) (22)	Ignored
Hooking in convex strokes (6a)	Inherited (from source strokes) (13)	Ignored	Ignored
Rightward hooking on vertical strokes (6b)	Ignored	Ignored	Ignored
Hooking on horizontal strokes (6c)	Ignored	Mostly inherited (from source characters) (24)	Ignored

Leftward hooking on vertical strokes (6d, e)	Ignored	Inherited (from source characters) (25)	Ignored
Diagonalization of bottom horizontal stroke in constituent to left of another (7a, b, d)	Not applicable	Not applicable	Borrowed (32a-e)
Dotting/shortening of lower right diagonal stroke in constituent to left of another (7c)	Not applicable	Not applicable	Borrowed (32f, g)

One may speculate on possible social and functional reasons as to why none of the derivative systems borrowed Chinese character phonology in full: perhaps their writing cultures wanted to distinguish themselves from that of China, perhaps sound-representing katakana and hangul worked better when visually distinct from logographic Chinese characters in mixed-systems text, and certainly hangul originated in an utterly non-Chinese geometric logic. Factors like these have been addressed in many previous studies on script borrowing (Tranter 2001, Unseth 2005, Miyake 2017a, Handel 2019); for example, Handel (2019) explains how the typological features of spoken Japanese and Korean encouraged the development of non-logographic scripts in the first place.

Regarding the less-studied formal features shared by these four systems, some can be ascribed to universal tendencies, like the preference for cardinal axes, while others seem to have involved the genuine borrowing of formal regularities, and not merely script elements like characters or strokes. The likelihood that a regularity was borrowed seems to have depended in part on its productivity within Chinese script: enlargement is highly regular and was readily borrowed, the more exception-prone left-edge curving tended to be modified or ignored, and the highly lexicalized hooking was never borrowed as a productive process at all. Moreover, even when Chinese regularities were modified in the adopting systems, these systems still have script phonologies of their own: katakana non-left-edge curving and hangul left-edge enlargement may differ from Chinese but they are still internally consistent. The borrowing of script form regularities also applied at more than one level of representation. At the most superficial level (script phonetics) was the influence that the lexically non-distinctive off-horizontal stroke in Chinese <乜> had on the similarly non-distinctive off-horizontal stroke of katakana <ヤ> (24a). More abstract (script phonology) was the application of bottommost enlargement in vertical strokes in hangul <ㅗ> and <ㅜ>, despite the lack of direct Chinese models. Still more abstract was the sensitivity of left-edge curving to constituent boundaries (script morphology), borrowed from Chinese into Tangut script.

There are many more Sinoform scripts than these, of course. Some seem to have adopted Chinese character phonology wholesale, creating new constituents by combining and modifying standard Chinese strokes in standard ways, as in the semi-syllabic Khitan small script (Wu and Janhunen 2014), or in logographic scripts that also adapted Chinese character morphology (Handel 2019): Khitan large script, Jurchen script, Korean hanja, Japanese kanji, Vietnamese Chữ Nôm (see also Myers 2019, section 2.3.2), and Zhuang script. Other Sinoform scripts have such a different stroke inventory that standard Chinese regularities are generally inapplicable, such as the semi-syllabic Japanese hiragana and syllabic Sinitic Nüshu

script (Zhao 1998, Van Esch 2017), as well as the Yi script (Shi 1996), both in the classical logographic script (which also borrowed aspects of character morphology) and in the modern syllabic script. Even a quick glance, however, suggests that each of these systems has developed a coherent script phonology of its own.

There is no reason not to extend the analogy between areal script form patterns and areal sound patterns beyond the Sinosphere, although in more pronunciation-transparent systems (like those based on the Roman alphabet) we must also take into account interactions with spoken phonology. The keys are to distinguish the effects of borrowing from those of universals and inheritance, to supplement the field's current focus on the social and functional context of script borrowing with a search for psychologically plausible and relatively explicit formal principles underlying each script's "look", and above all, to recognize that just as phonology is not mere physics, even outer form has a rich inner life.

References

Ahn, Jinsu (2017). The dynamism of Hangeul's vertical strokes and the flow of its lines of writing. Visible Language 51 (3-1): 56-73.

Aikhenvald, Alexandra Y. (2001). Areal diffusion, genetic inheritance, and problems of subgrouping: A North Arawak case study. In Alexandra Y. Aikhenvald & R. M. W. Dixon (eds.), Areal diffusion and genetic inheritance: Problems in comparative linguistics, 167-194. Oxford: Oxford University Press.

Aronoff, Mark (1994). Spelling as culture. In W.C. Watt (ed.), Writing systems and cognition, 67-86. Springer: Berlin.

Beckman, Mary E. & Jan Edwards (1990). Lengthening and shortening and the nature of prosodic constituency. In John Kingston & Mary E. Beckman (eds.), Papers in laboratory phonology I: Between the grammar and the physics of speech, 152-178. Cambridge: Cambridge University Press.

Berent, Iris. (2013). The phonological mind. Trends in Cognitive Sciences 17 (7): 319-327.

Blevins, Juliette. (2017). Areal sound patterns: From perceptual magnets to stone soup. In Raymond Hickey (ed.), The Cambridge handbook of areal linguistics, 88-121. Cambridge: Cambridge University Press.

Brentari, Diane. (2011). Sign language phonology. In John Goldsmith, Jason Riggle & Alan C. L. Yu (eds.), The handbook of phonological theory, second edition, 691–721. Hoboken, NJ: Wiley-Blackwell.

Changizi, Mark A., Qiong Zhang, Hao Ye, & Shinsuke Shimojo. (2006). The structures of letters and symbols throughout human history are selected to match those found in objects in natural scenes. The American Naturalist 167 (5): E117-E139.

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

Dimmendaal, Gerrit J. (2001). Areal diffusion versus genetic inheritance: An African perspective. In Alexandra Y. Aikhenvald & R. M. W. Dixon (eds.), Areal diffusion and genetic inheritance: Problems in comparative linguistics, 358-392. Oxford: Oxford University Press.

Evertz, Martin (2018). Visual prosody: The graphematic foot in English and German. Berlin: Walter de Gruyter.

Fuhrhop, Nanna, Franziska Buchmann & Kristian Berg (2011). The length hierarchy and the graphematic syllable: Evidence from German and English. Written Language & Literacy 14 (2): 275-292.

Galambos, Imre (2015). Translating Chinese tradition and teaching Tangut culture: Manuscripts and printed books from Khara-Khoto. Berlin: Walter de Gruyter.

Gelb, Ignace J. (1963). A study of writing (revised edition). Chicago: University of Chicago Press.
Gong, Hwang-Cherng (1982). Chinese elements in the Tangut script. Bulletin of the Institute of History and Philology, Academia Sinica 53 (1): 167-187.
Handel, Zev (2019). Sinography: The borrowing and adaptation of the Chinese script. Leiden: Brill.
Hansell, Mark (2002). Functional answers to structural problems in thinking about writing. In Mary S. Erbaugh (ed.), Difficult characters: Interdisciplinary studies of Chinese and Japanese writing, 124-264. Columbus, Ohio: Ohio State University.
Hargus, Sharon & Ellen M. Kaisse (eds.) (1993). Studies in lexical phonology. Cambridge, MA: Academic Press.
Jakobson, Roman (1931). Über die phonologischen Sprachbünde. Travaux du Cercle Linguistique de Prague, 4, 234-240.
Kim, Zong-Su (2005). The history and future of Hangeul: Korea's indigenous script. (English translation by Ross King; original publication 1990). Folkestone, Kent: Global International, Ltd.
King, Ross (1996). Korean writing. In Peter T. Daniels & William Bright (eds.) The world's writing systems, 218-227. Oxford: Oxford University Press.
Kohrt, Manfred (1986). The term 'grapheme' in the history and theory of linguistics. In Gerhard Augst (ed.), New trends in graphemics and orthography, 80-96. Berlin: Walter de Gruyter.
Kychanov, E. I. (1996). Tangut. In Peter T. Daniels & William Bright (eds.), The world's writing systems, 228-230, 237. Oxford: Oxford University Press.
Lee, Chia-Ying, Jie-Li Tsai, Yu-Chin Chiu, Ovid J. L. Tzeng & Daisy L. Hung (2006). The early extraction of sublexical phonology in reading Chinese pseudocharacters: An event-related potentials study. Language and Linguistics 7 (3): 619-636.
Lee, Sang-Oak (2009). The Korean alphabet: an optimal featural system with graphical ingenuity. Written Language & Literacy 12 (2): 202-212.
Li, Fanwen (1997). Xià-Hàn Zìdiǎn [Tangut-Chinese dictionary], first edition. Beijing: China Society Science Publishing House.
Li, Fanwen (2008). Xià-Hàn Zìdiǎn [Tangut-Chinese dictionary], second edition. Beijing: China Society Science Publishing House.
McCawley, James D. (1994). Some graphotactic constraints. In W.C. Watt (ed.), Writing systems and cognition: Perspectives from psychology, physiology, linguistics, and semiotics, 115-127. Berlin: Springer.
Meletis, Dimitrios. (2018). What is natural in writing? Prolegomena to a Natural Grapholinguistics. Written Language & Literacy 21 (1) 52-88.
Meletis, Dimitrios. (2019). The grapheme as a universal basic unit of writing. Writing Systems Research 11 (1) 26-49.
Miyake, Marc (2017a). Minority languages: Traditional and modern writing systems. In Rint Sybesma, Wolfgang Behr, Yueguo Gu, Zev Handel, C.-T. James Huang & James Myers (eds.), Encyclopedia of Chinese language and linguistics, vol. III, 35-40. Leiden, Netherlands: Brill.
Miyake, Marc (2017b). Tangut language. In Rint Sybesma, Wolfgang Behr, Yueguo Gu, Zev Handel, C.-T. James Huang & James Myers (eds.), Encyclopedia of Chinese language and linguistics, vol. IV, 267-274. Leiden, Netherlands: Brill.
Morin, Olivier (2018). Spontaneous emergence of legibility in writing systems: The case of orientation anisotropy. Cognitive Science 42 (2): 664-677.
Myers, James (2019). The grammar of Chinese characters: Productive knowledge of formal patterns in an orthographic system. Abingdon: Routledge

Ninio, Anat & Amia Lieblich (1976). The grammar of action: 'Phrase structure' in children's copying. Child Development 47(3): 846-850.
Pacton, Sébastien, Pierre Perruchet, Michel Fayol & Axel Cleeremans (2001). Implicit learning out of the lab: The case of orthographic regularities. Journal of Experimental Psychology: General 130 (3): 401-426.
Peng, Xuanwei (2017). Stroke systems in Chinese characters: A systemic functional perspective on simplified regular script. Semiotica 2017 (218): 1-19.
Primus, Beatrice (2004). A featural analysis of the Modern Roman Alphabet. Written Language & Literacy 7(2): 235–274.
Qiu, Xigui (2000). Chinese writing [Compiled and translated from Chinese by Gilbert L. Mattos & Jerry Norman]. Berkeley: The Society for the Study of Early China and The Institute of East Asian Studies, University of California, Berkeley.
Sandler, Wendy (1993). A sonority cycle in American Sign Language. Phonology 10 (2): 243-279.
Seidenberg, Mark S., David C. Plaut, Alan S. Petersen, James L. McClelland & Ken McRae (1994). Nonword pronunciation and models of word recognition. Journal of Experimental Psychology: Human Perception and Performance 20 (6): 1177-1196.
Shi, Dingxu (1996). The Yi Script. In Peter T. Daniels & William Bright (eds.), The world's writing systems, 239-243. Oxford: Oxford University Press.
Song, Ki Joong (1998). The writing systems of Northeast Asia and the origin of the Korean alphabet, Han'gŭl. Seoul Journal of Korean Studies 11: 3-50.
Thomason, Sarah Grey & Terrence Kaufman (1988). Language contact, creolization, and genetic linguistics. Berkeley: University of California Press.
Tranter, Nicolas (2001). Script 'borrowing', cultural influence and the development of the written vernacular in East Asia. In Thomas E. McAuley (ed.), Language change in East Asia, 180-204. London: Routledge.
Tranter, Nicolas (2002). The 'ideal square' of logographic scripts and the structural similarities of Khitan script and han'gŭl. In Sang-oak Lee, Gregory K. Iverson, with Sang-cheol Ahn & Young-mee Yu Cho (eds.), Pathways into Korean language and culture: Essays in honor of Young-key Kim-Renaud, 503-522. Seoul: Pagijong Press.
Tsai, Kuan-Chan & Terezinha Nunes (2003). The role of character schema in learning novel Chinese characters. In Catherine McBride-Chang & Hsuan-Chih Chen (eds.), Reading development in Chinese children, 109-125. Westport, Connecticut: Praeger.
Tseng, Yuho. (1993). A history of Chinese calligraphy. Hong Kong: Chinese University Press.
Tsukishima, Hiroshi (1977). Dairokushō katakana no jigen [Chapter 6: The origin of katakana]. In Hiroshi Tsukishima, Kokugo no rekishi [History of Japanese], 81-108. Tokyo: University of Tokyo Press.
Tsukishima, Hiroshi (1997). Katakana no rekishi-teki kenkyū [Historical study of Japanese 'kata-kana' characters]. Nihon Gakushiin Kiyō [Transactions of the Japan Academy] 51 (3): 227-251.
Unseth, Peter (2005). Sociolinguistic parallels between choosing scripts and languages. Written Language & Literacy 8 (1): 19-42.
Van Esch (2017). Nǚshū (Women's script). In Rint Sybesma, Wolfgang Behr, Yueguo Gu, Zev Handel, C.-T. James Huang & James Myers (eds.), Encyclopedia of Chinese language and linguistics, vol. III, 262-267. Leiden: Brill.
Venezky, Richard L. (1967). English orthography: Its graphical structure and its relation to sound. Reading Research Quarterly 2 (3): 75-105.
Wang, Jason Chia-Sheng (1983). Toward a generative grammar of Chinese character structure and stroke order. University of Wisconsin-Madison Ph.D. thesis.

Watt, W.C. (1975). What is the proper characterization of the alphabet? I: Desiderata. Visible Language 9 (4): 293-327.
Watt, W.C. (2015). What is the proper characterization of the alphabet? VII: Sleight of hand. Semiotica 2015 (207): 65-88.
Whitman, John. (2011). The ubiquity of the gloss. Scripta 3: 95-121.
Wiedenhof, Jeroen (2017). Script reform (20th Century): History and analysis. In Rint Sybesma, Wolfgang Behr, Yueguo Gu, Zev Handel, C.-T. James Huang & James Myers (eds.), Encyclopedia of Chinese language and linguistics, vol. III, 671-676. Leiden: Brill.
Wiese, Richard. (2004). How to optimize orthography. Written Language & Literacy 7 (2): 305-331.
Wu, Yingzhe & Juha Janhunen (2014). New materials on the Khitan small script: A critical edition of Xiao Dilu & Yelü Xiangwen. Folkestone, Kent: Global Oriental.
Zhao, Liming (1998). Nüshu: Chinese women's characters. International Journal of the Sociology of Language 129: 127-137.