Quantifying Sound-Graphic Systematicity

Application to Multiple Phonographic Orthographies

Hana Jee · Monica Tamariz · Richard Shillcock

Abstract. Do letter-shapes predict in any way the canonical sounds they represent? Does the letter <a> in any sense visually predict its canonical pronunciation /æ/? We extended existing quantitative approaches to measuring systematicity between phonology and semantics. We quantified all pairwise visual distances between letters, using Hausdorff distance. We took the corresponding canonical pronunciations of the letters and quantified all pairwise distances between their feature-level representations, using edit distance and Euclidean distance. We defined letter-sound systematicity as a correlation between these two lists of distances. We confirmed Korean as the gold standard for letter-sound systematicity; it was designed in the 15th century to have exactly this characteristic. We found small but significant correlations in Arabic, Cyrillic, English, Finnish, Greek and Hebrew orthographies, with Courier New giving the most consistent correlations. Pitman's English shorthand and the Shavian alphabet also showed robust systematicity, and baseline fictitious orthographies showed no systematicity, validating our approach.

1. Background

It is a natural question whether certain parts of a letter or character are topologically related to its meaning or sound. This idea was in fact realized as hieroglyphs or logographs whose written characters are visually

Hana Jee 0000-0001-6248-9786
Psychology, School of Philosophy Psychology and Language Sciences, The University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ, Scotland
E-mail: `hana.jee@ed.ac.uk`

Monica Tamariz 0000-0003-4688-1774
Psychology, School of Philosophy Psychology and Language Sciences, Heriot-Watt University
E-mail: `m.tamariz@hw.ac.uk`

Richard Shillcock 0000-0002-0616-3703
Psychology, School of Philosophy Psychology and Language Sciences, The University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ, Scotland
E-mail: `rcs@inf.ed.ac.uk`

Y. Haralambous (Ed.), *Grapholinguistics in the 21st Century 2020. Proceedings*
Grapholinguistics and Its Applications (ISSN: 2681-8566, e-ISSN: 2534-5192), Vol. 5.
Fluxus Editions, Brest, 2021, pp. 905–925. `https://doi.org/10.36824/2020-graf-jeeh`
ISBN: 978-2-9570549-7-8, e-ISBN: 978-2-9570549-9-2

iconic. For example, the Chinese character <人> 'man' changes its size and location—as in <从> 'to follow' or in <囚> 'to lock up'—maintaining its original meaning in the different contexts. This iconicity facilitates learning the orthography (Dingemanse et al., 2015). For example, the location of the dot distinguishes the meanings between <犬> 'a dog' and <太> 'huge'. It is hard to explain why it is not the other way around, until one knows the former character visually represents a dog wagging its tail.

Phonographs allow far more room for arbitrariness between letters and the corresponding sound unit, but even phonograph users have attempted to theorize about letter shapes in a similar manner: the Roman letter <A> represents a bull's horn upside down; <O> represents the mouth shape of /o:/; and as <S> looks like a snake, it naturally sounds /s/ (Robinson, 1995). The reason why these speculations remained as speculations is related to the problem of this sort of rationalization: there is no consistent theory to apply to all the letter shapes. Such explanations seem to be based on somewhat haphazard analogy.

What is the origin of this propensity to think that meaning inheres in unmotivated written symbols? Looking at how writing emerged may provide an answer. Visual representation started with describing concrete objects (ibid.), but this must have involved some larger semantic value than the object itself—an intention, for example. It is likely that the anonymous painter of the Great Black Bull in the Lascaux cave retrieved the impression of a bull when wishing for a successful hunt.

There are a few scenarios regarding the emergence of writing. One of them suggests that the emergence of agriculture required record-keeping. Instead of a hand-to-mouth lifestyle, people had to remember, for example, the amount of a harvest and the proportion of seed-corn (Schmandt-Besserat, 1989). Many researchers agree that the act of writing began for business purposes: as communities grew and cities were formed, larger scale trade appeared (Havelock, 1976; Robinson, 1995; Rogers, 2005). These proto-writings (Robinson, 1995) occurred in various media, like sticks with notches, clay tokens, and numerical tablets, implying the necessity of simpler, quicker recording. At the same time, administrative procedures, such as those involving tax and the distribution of the population, were required, as in the Sumerian capital, Uruk (Sampson, 1985). These all indicate that the first writing involved forms of numbers—abstract concepts but unmotivated logographs (Havelock, 1976; Robinson, 1995; Rogers, 2005; Schmandt-Besserat, 1989). The ruins of the Assyrian empire (the first millennium BC) showed that their writing did not resemble pictography any more (Robinson, 1995). They managed to establish an arbitrary connection between written symbols and their connotations. However, these symbols did not yet acquire the status of phonographs, not being connected to individual sound units.

The idea of phonographic symbols appeared only after the discovery of the 'Rebus principle' (ibid.; Rogers, 2005), where a single pictographic symbol could be connected to a sound value. The sound unit at this stage was not necessarily phonemic and more likely syllabic (Havelock, 1976; Robinson, 1995). The more specific, phonemic association required the ability to segment the continuous flow of vocal sounds and to realize the differences between air flows and articulatory obstructions (Havelock, 1976). The presence or absence of vibration in the larynx (voiced vs. voiceless) also had to be noticed.

Once an alphabet set is established, the constituents of the system need to be in balance between *efficiency* and *distinctiveness*: they should be easy to write and distinguished from one another. For extreme efficiency, all the letters might look the same but at the cost of distinction. At the other extreme, the letters might differ in shape size, colour and material as well as orientation. In fact, many phonographic orthographies in human history satisfy *coherent discrimination* among the constituents of the system by addition, subtraction, duplication, or orientation change. It is plausible to expect these scripts to have developed some sort of *systematicity*, in order to make the best use of the limited resources to facilitate acquisition and transmission of the orthography.

As an exclusively cultural heritage (Sampson, 1985), each writing script undergoes its own cultural evolution. Removing inefficiencies and not creating a new revolutionary feature, is the central role of cultural evolution (ibid.), a process that can be enhanced by repetitions and extensive communication. Multiple factors condition letter shapes. The nature of writing materials (Sirat, 1994) decides the angularity of letters: C versus <. The combination of writing materials and writing postures also affect the complexity of letters: compare a pen on a paper and a chisel on a clay tablet. Watt (2013) pointed out that letters tend to face the same direction; the facing direction is defined as the direction of ornaments and headings—for example, Arabic numbers mostly face left. Anecdotal evidence says that children often reverse, for instance, the letter B until they subconsciously understand that asymmetric English letters generally face rightwards. The direction of the script can be also changed for political and cultural reasons. People in conquered territories often had to adapt to a new writing custom. For example, Egyptians began to write from left to right when they accepted Christianity but later returned to write from right to left when Islam prevailed in the region in the 7C. The direction of script affects the direction of letters because moving backwards slows the pace of continuous writing, reducing efficiency. Hebrew seems to take longer to write because the overall script moves from right to left whereas horizontal strokes are written from left to right (Sirat, 1994). In Rome and Greece, the scripts were written successively from right to left, and then left to right, termed boustrophedon fashion or ox-turning. Asymmetric letters like B, E, N were frequently written in their mirror images to match the direction of the script. Some of the letters of the modern Roman alphabet therefore remained as their mirror images when the writing direction was stabilized from left to right (Sirat, 1994).

Watt (1979) introduced four hypothetical forces that affect letter shapes. *Homogenization* means the letters look more alike; *heterogenization* means they are distinguished from each other. For example, the uppercase letters <D>, <E>, <F>, <H> are visually homogeneous whereas their lowercase counterparts <d>, <e>, <f>, <h> are considerably heterogenized. *Facilitation* means the tendency for letters to be easy to produce. For instance, cursive movement minimizes direction shifts and hand movements, for greater writing speed (Sirat, 1994). Finally, *inertia*, is a conservative force to stabilize the system. These forces are more topological than kinetic. When they are in equilibrium, the orthography system stays the same, but when any of the first three forces gets stronger, letter shapes may change. Any change of letter shapes or introduction of a new letter occurs *in connection with* the other elements in the system, the other letter shapes and sounds (Brekle, 1994; Watt, 1979, Watt, 1994, Watt, 2013).

In this paper, we suggest a novel approach to investigating the assumed systematicity between letters and sounds. It is, however, not a symbolic logic in which "letter <g> has a definite feature of its sound /g/". It is rather a reflection of the system as a whole: "is letter <g> close to letter <k> as much as the sound /g/ is close to the sound /k/? How much do the distances among the phonemes correlate with the distances among the visual representations of those phonemes? To our knowledge, this is the first such quantitative demonstration of letter-sound systematicity across the whole alphabet.

We transferred this method from recent studies reporting systematicity between semantics and phonology (Dautriche, Mahowald, Gibson, and Piantadosi, 2017; Monaghan, Shillcock, Christiansen, and Kirby, 2014; Shillcock, Kirby, McDonald, and Brew, 2001; Tamariz, 2008). We explored the systematic relation between phonology and orthography in Arabic, Cyrillic, English, Finnish, Greek, Hebrew, and Korean.

2. Procedure

We measured all the pairwise visual distances between letters and the corresponding pairwise phonological distances between the canonical pronunciations of those letters, in the respective alphabets. The total pairwise distances in phonology or semantics are $N \times (N-1)/2$. We defined letter-sound systematicity as the correlation between the resulting two lists of distances, as in the Mantel Test (Mantel, 1967). The significance of the correlation between these two lists of pairwise distances was tested with a Monte-Carlo permutation test, as in the published literature on word-level systematicity. The whole process was conducted in Python 3.7.1.[1]

2.1. Phonological Distances

We encoded the phonemes of each language into feature vectors (cf. Farmer, Christiansen, and Monaghan, 2006) based on the International Phonetic Alphabet (IPA). The features consisted of place and manner of articulation. We marked 1 if a phoneme had the feature and 0 if it did not, and transformed each phoneme into a binary vector. For example, /b/ can be represented as [0,1,0,0,1,0,0]: palatal, labial, dental, throat, plosive, affricate, and fricative. The length of the vectors equalled the total number of phonological features of a language.

We measured the distances between two vectors as feature edit distance, which counts the number of features different between the two vectors, and as Euclidean distance, which measures the shortest geometric distance between two vectors. (Multiple distance metrics demonstrate the robustness of the results.) The more dissimilar two vectors are, the larger the values that are returned. For all phonological distance measures, we used `textdistance` 4.1.4 (Python 3.7.1)[2].

2.2. Orthographical Distances

We measured the distances between two letter images by Hausdorff distance (Huttenlocher, Klanderman, and Rucklidge, 1993). Hausdorff distance measures the difference between two images by first comparing each pixel of 'X' and 'Y' and then calculating Euclidean distance between the pixel from 'X' and the closest pixel from 'Y'. Being fundamentally asymmetric— the distance from 'X' to 'Y' is different from 'Y' to 'X'— the larger value is used by definition.

Because the letters were treated as images, different fonts returned different results. We examined various fonts available in Microsoft including serif, sans-serif, and cursive fonts: 29 fonts for Cyrillic, English, Finnish, and Greek (Table 1); 10 fonts for Arabic (Table 2); 13 fonts for Hebrew (Table 3); and 88 fonts for Korean (Appendix F). The letters were all centrally aligned with the default font setting and saved as an identically sized PNG image file. An implementation of Hausdorff distance in Python 3.7.1 converted these images into black and white raster graphics and returned numeric values as results.

1. The Python code are available from `https://github.com/HanaJee/hausdorff-distance-letters.git`.

2. Textdistance 4.1.4 imported from `https://pypi.org/project/textdistance/` in July 2019.

TABLE 1. Fonts examined for Cyrillic, English, Finnish, and Greek

Serif fonts	Book Antiqua, Cambria, Constantia, Courier New, Gabriola, Georgia, Lucida Console, Palatino Linotype, Times New Roman
Sans-serif fonts	Arial, Arial Black, Candara, Calibri, Calibri Light, Century Gothic, Comic Sans MS, Consolas, Corbel, Franklin Gothic Medium, Impact, Lucida Sans Unicode, Microsoft Sans Serif, Segoe UI Symbol, Tahoma, Trebuchet MS, Verdana
Cursive style	Lucida Handwriting, Segoe Print, Segoe Script

2.3. Samples

Arabic

A written Arabic alphabet (Arabic abjad) can have a maximum of four different forms: in the initial positions, in the middle of a word, in the final positions, and in the isolated forms (Erfani, 2005). We examined the isolated forms as they are the canonical letters that are first taught to children. Note that Arabic long vowels (<ا> /aː/, <و> /w/, and <ي> /j/) are included in the set of the alphabet, whereas short vowels (<◌ُ> /u/, <◌َ> /a/, and <◌ِ> /i/) are considered diacritics. We collected 28 Arabic letters and vectorized their corresponding phonemes based on 18 IPA features (Appendix A).

Cyrillic

Cyrillic script is used in many Eastern European countries, including Russia, but there are variations. Russian Cyrillic, for example, was reformed in the 18th century. Ukrainian, Bulgarian, Serbian, Macedonian, and Iranian script among others also look slightly different. We used the common Cyrillic letters and their phonemes (Appendix B). The letters <E>, <Ю> and <Я> were excluded because they are diphthongs (/jɛ/, /ju/ and /ja/, respectively), as was <Ь>, because it simply makes consonants softer and does not have any phonetic value. Accordingly, we made 25 phoneme vectors based on 20 IPA features (Appendix B).

English

As a deep orthography (Seymour, Aro, and Erskine, 2003), English letters are linked to more than one phoneme. We first constrained the sound of a letter according to the British phonics approach (Lloyd, Wernham, Jolly, and Stephen, 1998). Phonics teaches children the most frequent and canonical sound of the letter. We excluded <x> and <q> from the sample because the former is a polyphone /ks/ and the latter almost always co-occurs with <u>. In total, 24 letters were converted into feature-vectors, taken from Harm and Seidenberg (1999).

Finnish

Finnish script is the same as English except for three additional letters: <ä>, <ö>, and <å>, and the letters <k>, <p>, and <t> have tensed sounds, not aspirated. We included <q> because it is independently pronounced /k/. The 28 letters and 17 phonetic features are listed in Appendix C.

Greek

Greek uppercase letters are historically important in that they are closely related to ancient orthographies such as Phoenician. Lowercase letters have distinct forms from uppercase letters (Appendix D). The uppercase letter <Σ> (/s/) corresponds to two lowercase letters, which we included. We excluded <Ξ> and <Ψ>, as well as their corresponding lowercases <ξ> and <ψ>, because they are diphthongs: /ks/ and /ps/, respectively. We used 19 IPA features for the Greek phonemes.

Hebrew

As a consonantal orthography, written Hebrew for advanced readers does not indicate vowel values. The vowels are only written out for children and foreign learners until they get used to reading. We examined 33 consonants with 14 IPA phonetic features (Appendix E).

Korean

Hangeul, the Korean orthography, was artificially invented in the 15th century. It is well known for its one-to-one connection between letters and sounds, and for the fact that its letters were designed based on the shape of articulation. For example, <ㄱ> /g/ represents the tongue touching the soft palate. Korean phonology distinguishes between phonemes that are considered allophones by English speaker: /p/ in 'pie' and 'spy' are perceived as *aspirated* and *tensed*, respectively. Along with the *lenis* sound that shares the same articulation point without aspiration, these phonemes have visually systematic forms (e.g., <ㅂ> /b/ -<ㅍ> /p/ -<ㅃ> /p*/). Based on more cultural grounds, Korean written vowels are composed of three components: <·>, <ㅡ>, and < ㅣ >, which respectively represent the heaven, earth and human. In total, we examined 16 consonants and 10 monophthongs (Appendix F).

Other Orthographies

We additionally examined four ancient Semitic orthographies (Phoenician, Nabataean, Early Arabic, and Aramaic), two English substitute systems (*Pitman's shorthand* and the *Shavian alphabet*) and two fictitious orthographies (*Aurebesh* from *Star Wars* and *Klingon* from *Star Trek*) in terms of sound-letter systematicity. We expect if such a correlation is found in the modern conventional orthographies, it evolved over cultural time. We do not expect to observe any sound-letter systematicity in the fictitious systems that have not undergone natural selection in human culture. Finally, the artificially, consciously constructed letters in the Pitman's shorthand and Shavian alphabet may be expected to have a systematicity comparable to Korean orthography.

3. Results

General Results

For each orthography, we calculated systematicity as Pearson's r and confirmed the significance level with Monte-Carlo permutation tests. For each of the naturally occurring orthographies there were fonts for which significant systematicity obtained: for Korean 85 out of 88 fonts produced significant systematicity; for Finnish only 2 fonts out of 29 returned a significant systematicity. When a font exhibited significant systematicity, it was generally of the order of $r = 0.1 - 0.15$ (see Fig. 1; see below, also); similar letters tend to have similar sounds. Greek lower cases, in contrast, showed a negative correlation; similar letters tend to have distinct sounds.

Arabic

Table 2 indicates that Arabic letters tend to correlate with their sounds. Simplified Arabic consistently showed significant systematicity regardless of phonemic distance measure.

Cyrillic

Cyrillic upper and lower cases both correlated with the phonemes only in Courier New. The upper cases: $r = .14$, $p = .02$ when measured by Euclidean distance, $r = .18$, $p < .01$ when measured by feature edit distance. The lower cases: $r = .14$, $p = .02$ when measured by Euclidean distance and $r = .18$, $p < .001$ when measured by feature edit distance.

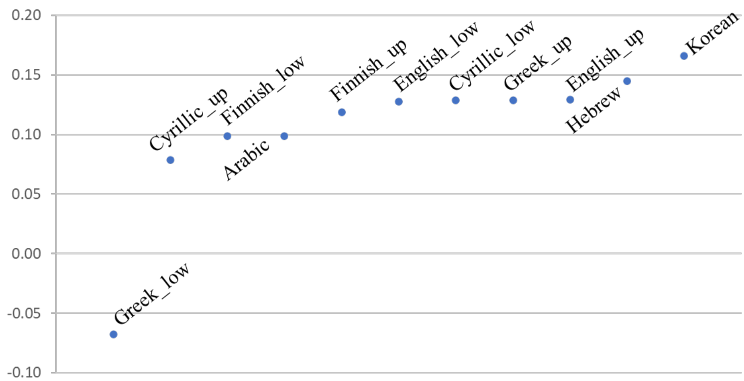

FIGURE 1. Letter-sound correlations of the conventional orthographies: we averaged the correlation coefficients from various fonts only when p-value $< .1$.

TABLE 2. The letter-sound correlations in 10 Arabic fonts. Note: * $p < .05$, ** $p < .01$, *** $p < .001$, $N = 378$, phonological distance $M = 1.46$, $SD = 0.27$ (Euclidean); $M = 2.46$, $SD = 1.43$ (feature edit); orthographical distance $M = 10.50$, $SD = 2.64$

		Euclidean distance			feature edit distance		
Font	Example	r	p-value		r	p-value	
Simplified Arabic	تشرفت بمقابلتك	0.15	< .001	***	0.1	0.05	*
Arial Black	تشرفت بمقابلتك	0.13	0.01	**	0.09	0.07	
Times New Roman	تشرفت بمقابلتك	0.13	0.01	**	0.09	0.07	
Arabic Typesetting	تشرفت بمقابلتك	0.12	0.02	*	0.02	0.74	
Traditional Arabic	تشرفت بمقابلتك	0.11	0.03	*	0.04	0.42	
Courier New	تـشـرفـت بـمـقـابـلـتـك	0.07	0.2		0.05	0.32	
Microsoft Sans Serif	تشرفت بمقابلتك	0.05	0.37		0.12	0.02	**
Segoe UI	تشرفت بمقابلتك	0.04	0.38		0.12	0.02	**
Andalus	تشرفت بمقابلتك	0.04	0.45		0.05	0.29	
Tahoma	تشرفت بمقابلتك	-0.01	0.83		0.11	0.03	**

English

For upper cases, Cambria consistently returned correlation: $r = .11$, $p = .07$ when measured by Euclidean distance, $r = .12$, $p = .04$ when measured by feature edit distance. Gabriola ($r = .12$, $p = .04$), Georgia ($r = .10$, $p = .08$), and Impact ($r = .15$, $p = .01$) additionally showed the result when measured by feature edit distance only. For lower cases, Franklin Gothic Medium ($r = .15$, $p = .01$), Arial Black ($r = .14$, $p = .02$), Verdana (r = .14, p = .02), Cambria (r = .13, p = .03), and Tahoma (r = .12, p = .04) returned the results only when measured by feature edit distance.

Finnish

Significant systematicity was found only in uppercase Courier New (r = .12, p = .02). For lower cases, Segoe Script (r = .10, p = .05 measured by Euclidean distance) and Trebuchet MS (r = .11, p = .04, measured by feature edit distance) returned coefficient above significance level.

Greek

For upper cases, Courier New consistently showed robust coefficient: r = .15, p = .02 measured by Euclidean distance, r = .16, p = .02 measured by feature edit distance. Book Antiqua (r = .13, p = .04) was also significant when measured by Euclidean distance. Although marginal, the lowercase Courier New returned the negative correlation (r = −.11, p = .09) when measured by Euclidean distance.

Hebrew

All 13 fonts returned highly significant correlation coefficients (Table 3). Some fonts returned letter-sound systematicity even higher than that of Korean orthography.

Korean

Almost all 88 Korean fonts returned significant letter-sound correlation, including a few representative fonts: 굴림: r = .24, $p < .001$; 바탕: r = .18, $p < .001$; 궁서: r = .30, $p < .001$; 맑은고딕: r = .18, $p < .001$. KCC 은영 returned the highest coefficient: r = .39, $p < .001$. We re-calculated the correlation excluding each letter to investigate which contributes the most to the whole correlation. Each letter seems to contribute approximately equally to the whole letter-sound correlation.

Other Orthographies

None of the four ancient orthographies returned significant systematicity; nor did the two fictitious orthographies. We conducted Monte-Carlo permutation tests for verification.

The two English substitute writing systems returned high positive letter-sound correlations: *Pitman's shorthand*, in which r = .35, $p < .001$; and the *Shavian alphabet*, r = .2, $p < .001$.

TABLE 3. The letter-sound correlation r in 13 Hebrew fonts (all $p - value < .001$). Note: $^{*} p < .05$, $^{**} p < .01$, $^{***} p < .001$, $N = 1035$, phonological distance $M = 1.91$, $SD = 0.43$ (Euclidean); $M = 3.14$, $SD = 1.79$ (feature edit); orthographical distance $M = 15.57$, $SD = 9.89$

Font	Example	Euclidean distance	Feature edit distance
Levenim MT	שלום	0.35	0.31
Narkisim	שלום	0.35	0.31
Miriam	שלום	0.34	0.31
Times New Roman	שלום	0.34	0.31
David	שלום	0.33	0.3
Lucida Sans Unicode	שלום	0.30	0.28
Gisha	שלום	0.27	0.27
Arial	שלום	0.25	0.26
Arial Black	שלום	0.25	0.26
Calibri Light	שלום	0.25	0.26
Microsoft Sans Serif	שלום	0.24	0.27
Courier New	שלום	0.21	0.23
Tahoma	שלום	0.18	0.24

4. Discussion

We explored the systematicity of letter-sound mapping over a number of orthographies, from a new perspective. Seven conventional orthographies, as well as two reformed spelling systems, demonstrated that letters to some extent correlate with their pronunciations. *Hangeul*, the systematically invented orthography with a sophisticated understanding of phonology, constitutes the highest benchmark of letter-sound correlation; other artificial orthographies, *Pitman's shorthand* and the *Shavian alphabet* returned a similarly high correlation. Letter-sound correlation increases when visually similar figures are linked to articulatorily similar phonemes (e.g., <ㅋ> /k/ - <ㄱ> /g/ or <ㅌ> /t/ - <ㄷ> /d/). This fact explains why Hebrew also demonstrated a high correlation. The visual difference of letter shapes efficiently categorise the place of articulation and distinguish voiceless from voiced sounds (e.g., /k/ - /x/ or /v/ - /b/ in Appendix 5). The systematicity of an orthography is enhanced when adding or subtracting a stroke or the orientation change of letter shapes occurs systematically with the corresponding phoneme pairs (e.g., voiced-voiceless).

Comparatively low coefficients of the orthographies with Roman alphabets (Cyrillic, English, Finnish, and Greek) may be attributable to their complicated history. They originated from Phoenician alphabets

(1000 BC), known as the first stable alphabetic script (Havelock, 1976; Robinson, 1995). It diverged to Hebrew and Greek, and the latter was borrowed by the Romans. The Roman alphabets spread through Europe and one of the lineages settled down as the English alphabets (Havelock, 1976; Robinson, 1995; Rogers, 2005). Some 3000 years of the history of this *Northwest Semitic Graeco-Roman-Etruscan alphabet* (Havelock, 1976) naturally allowed cultural intervention, sometimes organized (ibid.; Robinson, 1995; Rogers, 2005). For example, when Phoenician 22-consonant alphabets were accepted by Greeks, some phonetic values (mostly weak consonants) were changed to vowels. At the same time, three more vowels were added, resulting in 25 characters in total. Later, Runes, the Germanic alphabets entered Roman culture, influencing some of their letters: r, i, and b. Middle English went through the *Great English Vowel Shift*, as well as the distinction of upper cases from lower cases.

We expected the modern European alphabet systems to demonstrate stronger systematicity than the ancient orthographies. Four ancient orthographies did not show significant systematicity. The authenticity of the phonemes (and characters) recovered (Havelock, 1976; Robinson, 1995) may be not perfectly reliable.

In conclusion, the human brain is adept at taking advantage of any type of systematicity, from the level of the neural substrate to cross-modality processing (Bavelier and Neville, 2002; Spence, 2011). There are many demonstrations of audio-visual multisensory perception (Baier, Kleinschmidt, and Müller, 2006; Calvert, Brammer, et al., 1999; Calvert, Campbell, and Brammer, 2000; Calvert, Hansen, Iversen, and Brammer, 2001; Fiebelkorn, Foxe, and Molholm, 2010; Kriegstein and Giraud, 2006; Zangenehpour and Zatorre, 2010), some of which specifically focus on grapheme-phoneme relations (Raij, Uutela, and Hari, 2000; Atteveldt, Formisano, Goebel, and Blomert, 2004; Weissman, Warner, and Woldorff, 2004). Although the data generally imply that no area is exclusively related to reading, the human brain certainly has the wherewithal to take advantage of the type of systematicity we have demonstrated in the relation between letters and their canonical pronunciation. One potential process underlying the emergence of systematicity may be Zipf's principle of least effort (Zipf, 1949), whereby least effort in pronunciation travels with least effort in writing a character, with these processes conditioning the pairwise distances within phonological and visual space.

References

Atteveldt, N. van et al. (2004). "Integration of letters and speech sounds in the human brain." In: *Neuron* 43.2, pp. 271–282.

Baier, B., A. Kleinschmidt, and N.G. Müller (2006). "Cross-modal processing in early visual and auditory cortices depends on expected statistical relationship of multisensory information." In: *Journal of Neuroscience* 26.47, pp. 12260–12265.
Bavelier, D. and H.J. Neville (2002). "Cross-modal plasticity: where and how?" In: *Nature Reviews Neuroscience* 3.6, pp. 443–452.
Brekle, H.E. (1994). "Some thoughts on a historico-genetic theory of the lettershapes of our alphabet." In: *Writing systems and cognition: Perspectives from psychology, physiology, linguistics, and semiotics*. Ed. by W. Watt. Berlin: Springer.
Calvert, G.A., M.J. Brammer, et al. (1999). "Response amplification in sensory-specific cortices during crossmodal binding." In: *Neuroreport* 10.12, pp. 2619–2623.
Calvert, G.A., R. Campbell, and M.J. Brammer (2000). "Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex." In: *Current Biology* 10.11, pp. 649–657.
Calvert, G.A. et al. (2001). "Detection of audio-visual integration sites in humans by application of electrophysiological criteria to the BOLD effect." In: *Neuroimage* 14.2, pp. 427–438.
Dautriche, I., E. Chemla, and A. Christophe (2016). "Word learning: Homophony and the distribution of learning exemplars." In: *Language Learning and Development* 12.3, pp. 231–251.
Dautriche, I. et al. (2017). "Wordform similarity increases with semantic similarity: An analysis of 100 languages." In: *Cognitive science* 41.8, pp. 2149–2169.
Dingemanse, M. et al. (2015). "Arbitrariness, iconicity, and systematicity in language." In: *Trends in Cognitive Sciences* 19.10, pp. 603–615.
Erfani, I.M. (2005). *Arabic for Beginners*. Ottawa: Laurier Books Limited.
Farmer, T.A., M.H. Christiansen, and P. Monaghan (2006). "Phonological typicality influences on-line sentence comprehension." In: *Proceedings of the National Academy of Sciences* 103.32, pp. 12203–12208.
Fiebelkorn, I.C., J.J. Foxe, and S. Molholm (2010). "Dual mechanisms for the cross-sensory spread of attention: how much do learned associations matter?" In: *Cerebral Cortex* 20.1, pp. 109–120.
Harm, M.W. and M.S. Seidenberg (1999). "Phonology, reading acquisition, and dyslexia: insights from connectionist models." In: *Psychological review* 106.3, pp. 491–528.
Havelock, E.A. (1976). *Origins of Western Literacy. Four Lectures delivered at the Ontario Institute for Studies in Education*. Toronto: Ontario Institute for Studies in Education.
Huttenlocher, D.P., G.A. Klanderman, and W.J. Rucklidge (1993). "Comparing images using the Hausdorff distance." In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 15.9, pp. 850–863.
Kriegstein, K. von and A.L. Giraud (2006). "Implicit multisensory associations influence voice recognition." In: *PLoS biology* 4.10.

Lloyd, S. et al. (1998). *The phonics handbook*. Chigwell: Jolly Learning.
Mantel, N. (1967). *The detection of disease clustering and a generalized regression approach. Cancer Research*.
Monaghan, P. et al. (2014). “How arbitrary is language?” In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 369.1651.
Raij, T., K. Uutela, and R. Hari (2000). “Audiovisual integration of letters in the human brain.” In: *Neuron* 28.2, pp. 617–625.
Robinson, A. (1995). *The story of writing*. London: Thames and Hudson.
Rogers, H. (2005). *Writing systems: A linguistic approach*. Oxford: Blackwell.
Sampson, G. (1985). *Writing systems*. London: Hutchinson.
Schmandt-Besserat, D. (1989). *Two precursors of writing: Plain and complex tokens. The origins of writing*.
Seymour, P.H., M. Aro, and J.M. Erskine (2003). “Collaboration with COST Action A8 Network.” In: *British Journal of psychology* 94.2, pp. 143–174.
Shillcock, R.C. et al. (2001). “Filled pauses and their status in the mental lexicon.” In: *Proc. 2001 Conf. of Disfluency in Spontaneous Speech*, pp. 53–56.
Sirat, C. (1994). “Handwriting and the writing hand.” In: *Writing systems and cognition: Perspectives from psychology, physiology, linguistics, and semiotics*. Ed. by W. Watt. Berlin: Springer.
Spence, C. (2011). “Crossmodal correspondences: A tutorial review. Attention, Perception, and.” In: *Psychophysics* 73.4, pp. 971–995.
Tamariz, M. (2008). “Exploring systematicity between phonological and context-cooccurrence representations of the mental lexicon.” In: *The Mental Lexicon* 3.2, pp. 259–278.
Watt, W.C. (1979). “Iconic equilibrium.” In: *Semiotica* 28.1–2, pp. 31–62.
——— (1994). *Curves as angles. In Writing systems and cognition: Perspectives from psychology, physiology, linguistics, and semiotics (Vol*. Ed. by W. Watt. Berlin: Springer.
——— (2013). *Writing systems and cognition: Perspectives from psychology, physiology, linguistics, and semiotics (Vol. 6)*. Berlin: Springer.
Weissman, D.H. van, L.M. Warner, and M.G. Woldorff (2004). “The neural mechanisms for minimizing cross-modal distraction.” In: *Journal of Neuroscience* 24.48, pp. 10941–10949.
Zangenehpour, S. and R.J. Zatorre (2010). “Crossmodal recruitment of primary visual cortex following brief exposure to bimodal audiovisual stimuli.” In: *Neuropsychologia* 48.2, pp. 591–600.
Zipf, G.K. (1949). *Human behavior and the principle of least-effort*. Reading: Addison-Wesley.

A. Arabic

TABLE 4. Arabic letters and their phonemes

letter	phoneme	letter	phoneme	letter	phoneme	letter	phoneme
ا	/a/	د	/d/	ض	/dˤ/	ك	/k/
ب	/b/	ذ	/ð/	ط	/tˤ/	ل	/l/
ت	/t/	ر	/r/	ظ	/ðˤ/	م	/m/
ث	/θ/	ز	/z/	ع	/ʕ/	ن	/n/
ج	/d͡ʒ/	س	/s/	غ	/ɣ/	ه	/h/
ح	/ħ/	ش	/ʃ/	ف	/f/	و	/w/
خ	/x/	ص	/sˤ/	ق	/q/	ي	/j/

TABLE 5. The features of Arabic phonemes

Letters	Phonemes	Place of Articulation									Manner of Articulation								Vowel quality
		Labial	Dental	Denti-alveolar		Palatal	Velar	Uvular	Pharyngeal	Glottal	Nasal	Stop		Fricative		Affricate	Trill	Approximant	
				Plain	Emphatic							Voiceless	Voiced	Voiceless	Voiced	Voiced			
ا	a	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
ب	b	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
ت	t	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
ث	θ	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
ج	d͡ʒ	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0
ح	ħ	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0
خ	x	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0
د	d	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
ذ	ð	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
ر	r	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
ز	z	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
س	s	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
ش	ʃ	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0
ص	sˤ	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0
ض	dˤ	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0
ط	tˤ	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0
ظ	ðˤ	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0
ع	ʕ	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0
غ	ɣ	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0
ف	f	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
ق	q	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0
ك	k	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0
ل	l	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
م	m	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
ن	n	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
ه	h	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0
و	w	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0
ي	j	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0

B. Cyrillic

TABLE 6. Cyrillic letters and their phonemes

Upper	Lower	Phoneme	Upper	Lower	Phoneme
А	а	/a/	О	о	/o/
Б	б	/b/	П	п	/p/
В	в	/v/	Р	р	/r/
Г	г	/g/	С	с	/s/
Д	д	/d/	Т	т	/t/
Ж	ж	/ʒ/	У	у	/u/
З	з	/z/	Ф	ф	/f/
И	и	/i/	Х	х	/x/
Й	й	/j/	Ц	ц	/ts/
К	к	/k/	Ч	ч	/tʃ/
Л	л	/l/	Ш	ш	/ʃ/
М	м	/m/	Щ	щ	/ʃt/
Н	н	/n/			

TABLE 7. The features of Cyrillic phonemes

			Place of Articulation							Manner of Articulation						Vowel Qualities						
Upper letters	Lower letters	Phonemes	Voiced	Bilabial	Labio-dental	Alveolar	Palatal	Velar	Post-alveolar	Nasal	Plosive	Affricate	Fricative	Trill	Approximant	Close	Mid	Open	Front	Central	Back	Roundness
А	а	a	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0
Б	б	b	1	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
В	в	v	1	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
Г	г	g	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0
Д	д	d	1	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
Ж	ж	ʒ	1	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0
З	з	z	1	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
И	и	i	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0
Й	й	j	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
К	к	k	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0
Л	л	l	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
М	м	m	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Н	н	n	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
О	о	o	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	1
П	п	p	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
Р	р	r	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
С	с	s	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
Т	т	t	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
У	у	u	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1
Ф	ф	f	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
Х	х	x	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0
Ц	ц	ts	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
Ч	ч	tʃ	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0
Ш	ш	ʃ	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0
Щ	щ	ʃt	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0

C. Finnish

TABLE 8. Finnish letters and their phonemic features

		Place of Articulation						Manner of Articulation					Vowel Quality					
Letter	Phoneme	Voiced	Labial	Alveolar	Palatal	Velar	Glottal	Nasal	Plosive	Fricative	Approximant	Trill	Close	Mid	Open	Front	Back	Roundness
a	ɑ	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0
b	b	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
c	s	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0
d	d	1	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0
e	e	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0
f	f	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
g	g	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0
h	h	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0
i	i	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0
j	j	1	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0
k	k	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0
l	l	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0
m	m	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
n	n	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0
o	o	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1
p	p	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
q	k	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0
r	r	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0
s	s	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0
t	t	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0
u	u	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0
v	ʋ	1	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
w	ʋ	1	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
y	y	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	1
z	z	1	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0
ä	æ	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0
ö	ø	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	1
å	o	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1

D. Greek

Table 9. Greek letters and their phonemes

Upper	Lower	Phoneme	Upper	Lower	Phoneme
Α	α	/a/	Ν	ν	/n/
Β	β	/v/	Ο	ο	/o/
Γ	γ	/γ/	Π	π	/p/
Δ	δ	/ð/	Ρ	ρ	/r/
Ε	ε	/e/	Σ	σ	/s/
Ζ	ζ	/z/	Σ	ς	/s/
Η	η	/i/	Τ	τ	/t/
Θ	θ	/θ/	Υ	υ	/i/
Ι	ι	/i/	Φ	φ	/f/
Κ	κ	/k/	Χ	χ	/x/
Λ	λ	/l/	Ω	ω	/o/
Μ	μ	/m/			

Table 10. The features of Greek phonemes

			Place of Articulation							Manner of Articulation				Vowel Qualities							
Upper letter	Lower letter	Phonemes	Voiced	Bilabial	Labio-dental	Alveolar	Velar	Dental	Nasal	Plosive	Fricative	Lateral approximant	Trill	Close	Close-mid	Mid-back	Open	Central	Front	Back	Roundness
Α	α	a	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0
Β	β	v	1	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
Γ	γ	γ	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0
Δ	δ	ð	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0
Ε	ε	e	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0
Ζ	ζ	z	1	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
Η	η	i	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0
Θ	θ	θ	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0
Ι	ι	i	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0
Κ	κ	k	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0
Λ	λ	l	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
Μ	μ	m	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Ν	ν	n	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Ο	ο	o	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1
Π	π	p	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
Ρ	ρ	r	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
Σ	σ	s	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
Σ	ς	s	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
Τ	τ	t	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
Υ	υ	i	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0
Φ	φ	f	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
Χ	χ	x	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0
Ω	ω	o	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1

E. Hebrew

Table 11. Hebrew letters and their phonemes

Letter	Phoneme	Letter	Phoneme	Letter	Phoneme
א	empty	כּ	/k/	פּ	/p/
ב	/v/	כ	/x/	פ	/f/
בּ	/b/	ךּ	/k/	ף	/f/
ג	/g/	ך	/x/	צ	/ts/
ד	/d/	ל	/l/	ץ	/ts/
ה	/h/	מ	/m/	ק	/k/
ו	/v/	ם	/m/	ר	/r/
ז	/z/	נ	/n/	שׁ	/sh/
ח	/x/	ן	/n/	שׂ	/s/
ט	/t/	ס	/s/	תּ	/t/
י	/j/	ע	emp	ת	/t/

Table 12. The features of Hebrew phonemes

		Place of Articulation								Manner of Articulation					
Letter	Phoneme	Voiced	Bilabial	Labio-dental	Alveolar	Palatal	Velar	Glottal	Post-alveolar	Nasal	Plosive	Affricate	Fricative	Lateral approx.	Approximant
א	-	0	0	0	0	0	0	0	0	0	0	0	0	0	0
בּ	b	1	1	0	0	0	0	0	0	0	1	0	0	0	0
ב	v	1	0	1	0	0	0	0	0	0	0	0	1	0	0
ג	g	1	0	0	0	0	1	0	0	0	1	0	0	0	0
ד	d	1	0	0	1	0	0	0	0	0	1	0	0	0	0
ה	h	0	0	0	0	0	0	1	0	0	0	0	1	0	0
ו	v	1	0	1	0	0	0	0	0	0	0	0	1	0	0
ז	z	1	0	0	1	0	0	0	0	0	0	0	1	0	0
ח	x	0	0	0	0	0	1	0	0	0	0	0	1	0	0
ט	t	0	0	0	1	0	0	0	0	0	1	0	0	0	0
י	j	1	0	0	0	1	0	0	0	0	0	0	0	0	1
כּ	k	0	0	0	0	0	1	0	0	0	1	0	0	0	0
כ	x	0	0	0	0	0	1	0	0	0	0	0	1	0	0
ךּ	k	0	0	0	0	0	1	0	0	0	1	0	0	0	0
ך	x	0	0	0	0	0	1	0	0	0	0	0	1	0	0
ל	l	0	0	0	1	0	0	0	0	0	0	0	0	1	0
מ	m	0	1	0	0	0	0	0	0	1	0	0	0	0	0
ם	m	0	1	0	0	0	0	0	0	1	0	0	0	0	0
נ	n	0	0	0	1	0	0	0	0	1	0	0	0	0	0
ן	n	0	0	0	1	0	0	0	0	1	0	0	0	0	0
ס	s	0	0	0	1	0	0	0	0	0	0	0	1	0	0
ע	-	0	0	0	0	0	0	0	0	0	0	0	0	0	0
פּ	p	0	1	0	0	0	0	0	0	0	1	0	0	0	0
פ	f	0	0	1	0	0	0	0	0	0	0	0	1	0	0
ף	f	0	0	1	0	0	0	0	0	0	0	0	1	0	0
צ	ts	0	0	0	1	0	0	0	0	0	0	1	0	0	0
ץ	ts	0	0	0	1	0	0	0	0	0	0	1	0	0	0
ק	k	0	0	0	0	0	1	0	0	0	1	0	0	0	0
ר	ɣ	1	0	0	0	0	1	0	0	0	0	0	1	0	0
שׁ	ʃ	0	0	0	0	0	0	0	1	0	0	0	1	0	0
שׂ	s	0	0	0	1	0	0	0	0	0	0	0	1	0	0
תּ	t	0	0	0	1	0	0	0	0	0	1	0	0	0	0
ת	t	0	0	0	1	0	0	0	0	0	1	0	0	0	0

F. Korean

TABLE 13. Visually systematic Korean consonants

Voiced		Voiceless		Tensed	
ㄱ	/ g /	ㅋ	/ k /	ㄲ	/ k͈ /
ㄷ	/ d /	ㅌ	/ t /	ㄸ	/ t͈ /
ㅂ	/ b /	ㅍ	/ p /	ㅃ	/ p͈ /
ㅅ	/ s /			ㅆ	/ s͈ /
ㅈ	/ dʒ /	ㅊ	/ tʃ /	ㅉ	/ t͈ʃ /
ㅇ	/ ŋ /	ㅎ	/ h /		

TABLE 14. Korean mono-thongs included in the study

Mono-thongs	
ㅏ	/ a /
ㅓ	/ ʌ /
ㅗ	/ o /
ㅜ	/ u /
ㅐ	/ e /
ㅔ	/ ɛ /
ㅚ	/ ø /
ㅟ	/ y /
ㅡ	/ ɰ /
ㅣ	/ i /

Table 15. 88 Korean fonts examined

굴림	꽃길	FB이철수80목판M
돋움	개미똥구멍	FB이철수90목판TM
바탕	신과장	FB이철수90목판M
궁서	제주한라산체	FB이철수2000목판TM
맑은고딕	제주고딕체	FB이철수2001목판M
나눔고딕	제주명조체	FB이철수2001목판TM
나눔명조	부산체	Yoon다정
나눔손글씨붓체	고양체	Yoon민준
나눔손글씨 펜체	고양일산체	Yoon세희
나눔바른고딕	오성과한음체	Yoon아혜
나눔바른펜	막걸리체	Yoon지영
나눔스퀘어	전라북도체	Yoon지희
나눔스퀘어라운드	푸른전남체	Yoon형오
Noto Sans CJK KR	KoPub돋움체	Yoon홍수
Noto Serif CJK KR	KoPub바탕체	김남윤
도현체	다온청년고딕	이현지
주아체	EBS주시경체	윤태민
한나는11살체	EBS훈민정음	한둥근체 돋움
간이벽온방	EBS훈민정음새론체	한둥근체 바탕
대한민국독도체	KBIZ한마음고딕	KCC은영
법정체	KBIZ한마음명조	KCC김훈
대한체	도서관체	한글누리
월인석보체	호국체	기랑해랑체
고도체	간이벽온방	
아리따돋움	이롭게바탕체	
HS봄바람체2.0	tvN즐거운이야기체	
HS가을생각체	티몬몬소리체	
HS겨울눈꽃체	빙그레체	
HS두꺼비체	스웨거체	
가비아솔미체	한겨레결체	
가비아납작블럭체	조선일보명조체	
미생체	동그라미재단	
신비는일곱살	FB이철수80목판TM	