Systematicity in language and the fast and slow creation of writing systems: Understanding two types of non-arbitrary relations between orthographic characters and their canonical pronunciation

Hana Jee*

Monica Tamariz†

Richard Shillcock*

*Psychology, The University of Edinburgh

†Psychology, Heriot-Watt University

Author Note

Email: hana.jee@ed.ac.uk

Correspondence concerning this article should be addressed to Hana Jee, Psychology, School of Philosophy Psychology and Language Sciences, The University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ, Scotland.

ORCID: 0000-0001-6248-9786

Abstract

Words that sound similar tend to have similar meanings, at a distributed, sub-symbolic level (Monaghan et al., 2014). We extend this paradigm for measuring systematicity to letters and their canonical pronunciations. We confirm that orthographies that were consciously constructed to be systematic (Korean and two shorthand writing systems) yield significant correlations between visual distances between characters and the corresponding phonological distances between canonical pronunciations. We then extend the approach to Arabic, Hebrew, and English and show that letters that look similar tend to sound similar in their canonical pronunciations. We indicate some of the implications for education, and for understanding typical and atypical reading. By using different visual distance metrics we distinguish between *symbol-based* (Korean, shorthand) and *effort-based* (Arabic, Hebrew, English) *grapho-phonemic systematicity*. We reinterpret existing demonstrations of phono-semantic systematicity in terms of cognitive effort.

Keywords: systematicity; pronunciation; orthography; phonics

Statement of Relevance

Monaghan et al.'s (2014) "How arbitrary is language" has attracted wide interest across language- and cognition-oriented disciplines (239 citations, to date); they show that words that sound similar tend to have similar meanings. In our paper, we use the same logic of quantifying systematicity, but apply it to letters of the alphabet and their canonical pronunciations, for the first time. We show that individual letters that look similar tend to be pronounced in similar ways. We quantify 'symbol-based systematicity' in Korean and in shorthand systems, and for different fonts. Crucially, we discover 'effort-based systematicity' in Arabic, Hebrew, and English. We indicate the importance for Education (as in the phonics approach to teaching reading) and for understanding skilled and atypical (dyslexic) reading. Critically, we reassess the concept of systematicity (i.e. non-arbitrariness) in language in terms of human effort, and re-evaluate existing demonstrations of systematicity.

Non-arbitrariness between orthographic characters and their canonical phonology:

The nature of systematicity in language

Spoken language is characterized by systematic subdomains such as phonology, morphology, syntax, semantics and pragmatics. Their systematicity is driven by informational requirements, but also reflects the robust redundancy necessary in the vulnerable medium of speech. Cognitive scientists increasingly see perception and cognition in terms of 'predictive processing': active prediction drives language processing, compared with the traditional view of listeners/readers passively accumulating data prior to parsing and interpretation. Maximizing systematicity in language use maximizes predictability, facilitating fast, incremental processing.

Researchers have increasingly identified systematic relations *between* linguistic domains. For instance, words that sound similar tend to have similar meanings (Blasi, Wichmann, Hammarstrom, Stadler, & Christiansen, 2016; Dautriche, Mahowald, Gibson & Piantadosi, 2017; Monaghan, Shillcock, Christiansen, & Kirby, 2014; Shillcock, Kirby, McDonald, & Brew, 2001; Tamariz, 2008). This small but significant relationship is separate from morphological and/or etymological relatedness and more general than pockets of phonetic symbolism (see below). Monaghan et al. (2014) found the effect was largest in early-acquired words, suggesting systematicity is adaptive in early language learning and/or for words of central cultural utility. Tamariz (2008) found the same meaning-form correlation in Spanish, with consonants tending to contribute positively, vowels negatively.

Phonetic symbolism has long been recognized (Hinton, Nichols, & Ohala, 2006): word-initial /sn/ is associated with nasal meanings in English: *snore*, *sneeze*, *snout*, *snot*, *sniff*, *snuff*, etc. This systematicity may partly reflect words taking different historical routes into English. In addition, priming keeps rarer words in usage: *gleam*, *glow* and *glint* help the interpretation of *gloaming*. It may also reflect genuine iconicity operating through perceptuo-motor analogy (Dingemanse, Blasi, Lupyan, Christiansen, & Monahgan, 2015).

Phonological form systematically relates to syntactic category (Fitneva, Christiansen, & Monaghan, 2009; Kelly, 1992; Morgan & Demuth, 1996). For instance, English nouns are more likely to contain a nasal segment, English verbs a front vowel (Fitneva et al., 2009). Shi, Morgan and Allopenna (1998) show that schwa reliably indicates functor status across languages as different as English, Mandarin and Turkish.

However, one relation between domains has attracted less attention—that between written letters and their sounds: for instance, the relation between the visual form of the letter ‘a’ and its canonical English pronunciation /æ/, or between ‘c’ and /k/. Do letters that look similar have similar canonical pronunciations? Can we measure any such systematicity? As far as we are aware, there are no quantitative studies of grapho-phonemic systematicity comparable to those reviewed above for morpho-phonemic systematicity. Importantly, understanding systematicity between orthography and canonical pronunciation may have implications for education and literacy. In the phonics approach (cf. Castles, Rastle & Nation, 2018), children are taught to read and write by initially associating letters with their canonical pronunciations and then combining them into words: ‘c’, ‘a’ and ‘t’ spell ‘cat’.

Most skilled readers/writers think of Roman letters as arbitrary signs, such that ‘m’ or ‘b’ have no visual properties pertaining to their canonical pronunciations /m/ and /b/. In contrast, users of certain other orthographies are aware of conscious, organized intervention to enforce grapho-phonemic systematicity, as in the creation of Korean orthography, the simplification of traditional Chinese, the emergence of unpointed Hebrew (Frost, 1995), the development of pinyin for Standard Chinese, and the creation of shorthand (Pitman, 1845). Thus, we can distinguish between fast, consciously produced orthographies and others that have emerged by slower, cultural accretion.

Simons (2011) traces the development of the Roman alphabet back to antiquity. Algeo and Butcher (2013) and others document and motivate additions, deletions and changes in form. This *symbolic* analysis perhaps finds its greatest complexity in the application of Optimality Theory (Smolensky & Prince, 1993) to the relation between letters and their German pronunciation (Song & Wiese, 2010), showing that lowercase Roman letters have grapho-phonemic systematicity. In contrast, tapping *all* the relations between representations within a domain produces metrics that are perhaps better characterized as *distributed* or *subsymbolic* (cf. Blank, Meeden, & Marshall, 1992), as in the review of phono-semantic systematicity above.

Arabic and Hebrew script are well known for their long histories. Arabic script can be traced back to Assyrian period (7th century BC); Hebrew script diverged from Phoenician script by 9th century BC (Robinson, 1995). They both maintain the feature from Aramaic, the Semitic language that does not write vowels.

Korean hangeul was created in the 15th century by royal scholars and king Sejong the Great himself. It is a shallow orthography (Seymour, Aro, & Erskine, 2003): a letter and its canonical pronunciation have an exclusive one-to-one correspondence. Its internal design reflects phonetic features (Sampson, 1985); for example, ㄱ (/g/) represents the tongue touching the soft palate, and ㅅ (/s/) represents the airflow through the teeth. Vowels reflect a more cultural underpinning: the basic three vowels ㅡ, ㅣ, and ・ respectively symbolize the earth, a person and the sky. Their combination represents the harmony between people and nature.

Korean phonology distinguishes between phonemes that are considered allophones by English speakers: thus /p/ in ‘spy’, /t/ in ‘star’, and /k/ in ‘sky’ are each distinguished as tensed phonemes, leading to triads of consonants: lenis-aspirated-tensed. Because the letters represent features rather than just phonemes, Sampson (1984) has called hangeul a 'featural alphabet' in contrast to alphabetic orthographies. Critically, the consonant triads have consistent letter shapes. For the lenis ㄱ (/g/), for example, adding a stroke makes it aspirated ㅋ (/k/), and duplicating makes it tensed ㄲ (/kk/). This visual consistency prevails across Korean consonants: ㄷ-ㅌ-ㄸ, ㅂ-ㅍ-ㅃ, and ㅈ-ㅊ-ㅉ. Compared with the Roman alphabet, where the letter shapes for those phoneme pairs are dissimilar, the orthographic regularity of hangeul increases the systematic relations with the phonemes and was intended to facilitate learning.

Pitman shorthand (Pitman, 1845) and the Shavian alphabet (Robinson, 1995) were also intentionally created as systematic, unambiguous, orthographically shallow alphabets (Appendix 1). Their authors paired English phonemes with geometric shapes. These three consciously invented orthographies (Korean, Pitman shorthand, and Shavian alphabet) are ideal benchmarks for quantifying grapho-phonemic systematicity for canonical pronunciations.

We first established the conceptual validity of quantifying grapho-phonemic systematicity as the correlation between all the phonological distances between canonical pronunciations and all the corresponding visual distances between characters, for orthographies designed to be systematic. From this vantage point, we then assessed the status of grapho-phonemic systematicity in the Arabic, Hebrew and English alphabets. (N.B. We did not include diphthongs and digraphs.)

Procedure

Sample

We measured grapho-phonemic systematicity of four conventional orthographies (Arabic, English, Hebrew and Korean) and two artificial English orthographies (Pitman's shorthand and the Shavian alphabet). We examined 28 Arabic letters in their isolated forms, 27 Hebrew letters including the 5 variations, 28 Korean letters (18 consonants and 10 monophthong vowels), and 24 English letters (excluding ‘x’ and ‘q’ because the former is the polyphone /ks/ and the latter almost always co-occurs with ‘u’). We examined 24 consonantal letters from Pitman’s shorthand and the Shavian alphabet (Appendix 1).

Measuring phonological distances

There are multiple ways of measuring the phonological distance between the canonical pronunciations of individual letters (cf. Sanders & Chin, 2009). We opted for a feature-based phonological distance metric. Phonological features are the central way of theorizing about phonology. They can be used comparably for other languages. We defined each phoneme as a binary vector according to place and manner of articulation using Harm and Seidenberg’s (1999) classification (see Jee, 2021, for discussion of other schemas for phonological classification). We then calculated feature-edit distances, Euclidean distances, cosine distances and Jaccard distances between phoneme vectors (Monaghan et al., 2014; Jee, 2021)[1]. The distances represent the difference in the numbers and types of articulatory feature involved in pronouncing the phonemes.

As a deep orthography (Seymour, Aro, & Erskine, 2003), English frequently connects more than one phoneme to a letter (e.g. ‘c’ in ‘cite’ and ‘cat’) and more than one letter to a phoneme (e.g. /k/ in ‘cat’ or ‘kite’). Researchers have quantified the quasi-regular ways in which letters contribute to pronunciation in large English lexicons (Berndt, Reggia, & Mitchum, 1987; Gontijo, Gontijo, & Shillcock, 2003). We restricted the sound of a letter to the canonical one

[1] Feature edit distance counts the number of difference between vectors; Euclidean distance calculates the shortest geometric distance between vectors; cosine similarity measures the angle made by two vectors; and Jaccard similarity is the number of shared features divided by the size of the union between two vectors.

used in the British phonics approach (Lloyd, Wernham, Jolly & Stephen, 1998), which teaches the most frequent sound-letter relations (e.g. 'a' as /æ/).

Measuring orthographic distances

There are many ways of measuring visual distance between letters (cf. Gibson, Gibson, Pick & Osser, 1962; Kriegeskorte, Mur & Bandettini, 2008; see, also, developments in scalable vector graphics[2]). We chose pixel count, perimetric complexity (Pelli, Burns, Farell & Moore-Page, 2006) and Hausdorff distance (Huttenlocher, Klanderman, & Rucklidge, 1993). This suite of pixel-based metrics (a) ensured the robustness of any correlation with phonological distance; (b) could be equally applied to any orthography and to different fonts within the same writing system; and (c) were potentially sensitive to both feature-based (e.g. ascenders and descenders) and subsymbolic aspects of systematicity.

Pixel count was the simplest measure. Perimetric complexity is defined as ink area divided by perimeter. Hausdorff distance[3] measures similarity between forms. It treats letters or characters as images and quantifies the difference between them. It converts the images into black and white raster graphics. Every pixel in each of the two images in a corresponding location is compared. Given two sets of black pixels, X = {x1, … xn} and Y = {y1, … yn}, the directed Hausdorff distance is calculated as:

$$d_h(X,Y) = max(\underset{x \in X}{max}, \underset{y \in Y}{min}\ |x - y|)$$

where Euclidean distance measures the distance between two individual points, | x – y |. Being fundamentally asymmetrical (d_h (X,Y) ≠ d_h (Y,X)), the larger value between the two (max) was

[2] Scalable Vector Graphics (SVG) 2. *W3C. Retrieved 28 January 2021.*

[3] We have reported preliminary research using one instantiation of Hausdorff Distance as a measure of visual similarity of Korean characters at G21C 2020 Grapholinguistics in the 21st Century (Jee, Tamariz & Shillcock, 2020).

used[4]. Comparing the different metrics, we expected Hausdorff distance to be the most sensitive to systematic ('featural') typographic relations between characters, as found in Korean orthography (cf. Chang, Kim, Cha, & Kim, 2005) and pixel count to be least sensitive.

We analysed 22 fonts for Arabic, 310 fonts for English (both upper and lower cases), 39 fonts for Hebrew, and 153 fonts for Korean, available from Microsoft; Light, Regular, Bold, and Italic styles were considered as individual fonts. Each letter was presented in a 100x100 pixels format. Pixel count and perimetric complexity were calculated for each letter pair. For Hausdorff distance, the letters were saved as high-resolution image files in PNG format before the pairwise distances were calculated. The procedure was implemented in Python 3.7.[5]

Measuring systematicity

We defined grapho-phonemic systematicity as the correlation between all the pairwise distances between letter shapes and all the pairwise distances between their corresponding canonical pronunciations. For the 24 English letters, for instance, there were [(24×24)/2]-24, or 276, different orthographic distances and 276 corresponding phonological distances.

We calculated Pearson's *r* between the lists of distances. This correlation shows how much the level of similarity among letter-shapes corresponds to the level of similarity among phonemes. A positive correlation indicates that similar letter shapes tend to have similar sounds. A negative correlation indicates that similar letter shapes tend to have distinct sounds.

Following researchers in word-level form-meaning systematicity (see above), we verified the implications of the correlations by conducting Monte Carlo permutation tests (not reported

---

[4] We assessed the psychological validity of Hausdorff distance. Geyer (1977) reports a very sparse confusion matrix from judgements made on degraded isolated letters. The values of the matrix quantify the confusability of lowercase English letters. Our Hausdorff distances correlated significantly with both halves of the confusion matrix, in the expected direction ($r = -.294$, $p < .000$; $r = -.143$, $p = .018$).

[5] Code available from https://github.com/HanaJee/Pixel-Count/blob/main/measuring%20orthographic%20distances.txt.

here). We randomly paired all the orthographical distances and all the phonological distances 10,000 times and accumulated the correlation coefficients to see where the veridical correlation coefficient fell in the distribution.

We predicted significant grapho-phonemic systematicity for the consciously constructed orthographies of Korean and the two shorthand systems. We expected Hausdorff distance would be more sensitive to Korean's symbolic orthographic 'features'. We predicted the spontaneously developed orthographies would be less systematic than Korean even though, for instance, lowercase English letters show a relation between stops and ascenders ('b', 'd', 'k' and 't') and descenders ('g' and 'p'). The precise effects of different fonts are unpredictable.

Results

Grapho-phonemic systematicity of Korean, Pitman's and Shavian alphabets

Only Hausdorff distance returned significant grapho-phonemic systematicity for all three writing systems (Table 1; N.B. The z values from the Monte Carlo analyses aligned with the significance levels of the p values for the correlations). For Hausdorff distance, Korean showed significant systematicity for all but five fonts Table 2); visually similar letters have phonologically similar canonical pronunciations.

Table 1. Grapho-phonemic systematicity (Pearson's r) of Korean, Pitman's shorthand, and the Shavian alphabet

	Pixel count	Perimetric complexity	Hausdorff distance
Korean	0.07 (p = 0.14)	0.08 (p = 0.11)	0.24 (p < .0001)
Pitman's shorthand	0.03 (p = 0.60)	0.12 (p = 0.06)	0.18 (p = 0.004)
Shavian alphabet	0.10 (p = 0.10)	0.11 (p = 0.06)	0.19 (p = 0.001)

Note. The letter shapes were from https://unicode.org/L2/L2016/16196-pitman.pdf for Pitman's shorthand and https://www.unicode.org/charts/PDF/U10450.pdf for the Shavian alphabet.

Table 2. Grapho-phonemic systematicity of the top 5 and bottom 10 Korean fonts, calculated using Hausdorff distance

No	Font	Example	*r*	*p*	
1	KCC 은영	안녕하세요	0.39	0.00	***
2	HS 가을생각체	안녕하세요	0.37	0.00	***
3	KBIZ 한마음고딕	안녕하세요	0.37	0.00	***
4	HS 두꺼비체	안녕하세요	0.35	0.00	***
5	KBIZ 한마음명조	안녕하세요	0.35	0.00	***
		……			
79	이현지	안녕하세요	0.11	0.03	*
80	Yoon 다정	안녕하세요	0.1	0.05	*
81	Yoon 세희	안녕하세요	0.1	0.04	*
82	Yoon 형오	안녕하세요	0.1	0.04	*
83	Yoon 홍수	안녕하세요	0.1	0.05	*
84	한겨레결체	안녕하세요	0.08	0.13	
85	조선일보명조체	안녕하세요	0.08	0.09	
86	FB 이철수 2001 목판 M	안녕하세요	0.08	0.11	
87	FB 이철수 2001 목판 TM	안녕하세요	0.08	0.11	
88	신비는일곱살	안녕하세요	0.06	0.26	

Note. * $p < .05$, ** $p < .01$, *** $p < .001$. The phonological distances were measured by feature edit distance. For Tables 2 and 3, see Jee (2021) for results using other phonological distance metrics.

Grapho-phonemic systematicity of Arabic, Hebrew and English

The three spontaneously developed orthographies showed significant grapho-phonemic systematicity in multiple fonts. Table 3A and 3B show all the fonts examined and the grapho-phonemic systematicity in Arabic and Hebrew, respectively. For English, Table 3C shows the 24 most frequently used fonts out of 310 fonts examined. Similar letter shapes tend to have similar canonical pronunciations in these three languages, with pixel count revealing more robust systematicity than Hausdorff distance.

Table 3A. Grapho-phonemic systematicity of 31 Arabic fonts.

Font name	Example	Pixel count			Perimetric complexity			Hausdorff distance		
		r	p		r	p		r	p	
Aldhabi	اشكرك	0.02	0.64		0.01	0.84		0.04	0.49	
Andalus	اشكرك	0.10	0.05		0.03	0.62		0.08	0.11	
Arabic Typesetting	اشكرك	0.19	0.001	***	0.10	0.04	*	0.14	0.01	**
Arial	اشكرك	0.14	0.01	**	-0.01	0.81		0.08	0.10	
Arial Bold	**اشكرك**	0.14	0.001	***	0.00	0.95		0.05	0.32	
Calibri	اشكرك	0.19	0.001	***	0.07	0.15		0.15	0.001	***
Calibri Bold	**اشكرك**	0.19	0.001	***	0.08	0.14		0.15	0.001	***
Calibri Light	اشكرك	0.20	0.001	***	0.13	0.01	**	0.15	0.001	***
Courier New	اشكرك	0.16	0.001	***	0.02	0.66		0.04	0.44	
Courier New Bold	**اشكرك**	0.09	0.10		0.05	0.32		0.02	0.72	
Majalla	اشكرك	0.19	0.001	***	0.09	0.07		0.11	0.04	*
Majalla Bold	**اشكرك**	0.17	0.001	***	0.09	0.08		0.12	0.02	*
Microsoft Sans Serif	اشكرك	0.13	0.01	**	0.13	0.01	**	0.10	0.06	
Microsoft Uighur Bold	**اشكرك**	0.18	0.001	***	0.04	0.48		0.09	0.07	
Microsoft Uighur	اشكرك	0.18	0.001	***	0.07	0.19		0.10	0.05	
Segoe UI	اشكرك	0.10	0.06		0.06	0.28		0.05	0.33	
Segoe UI Bold	**اشكرك**	0.11	0.04	*	0.04	0.48		0.05	0.31	
Segoe UI Light	اشكرك	0.07	0.15		0.11	0.04	*	0.07	0.18	
Segoe UI Semilight	اشكرك	0.11	0.03	*	0.06	0.23		0.06	0.23	
Segoe UI Semibold	**اشكرك**	0.10	0.05	*	0.07	0.20		0.05	0.32	
Simplified Arabic Bold	**اشكرك**	0.11	0.03	*	0.00	0.98		0.09	0.08	
Simplified Arabic Fixed	اشكرك	0.21	0.001	***	0.03	0.60		0.11	0.03	*
Simplified Arabic	اشكرك	0.14	0.01	**	0.02	0.75		0.12	0.02	*
Tahoma	اشكرك	0.11	0.03	*	0.10	0.05	*	0.04	0.47	
Tahoma Bold	**اشكرك**	0.10	0.06		0.10	0.04	*	0.02	0.77	
Times New Roman	اشكرك	0.14	0.01	**	-0.01	0.83		0.09	0.10	
Times New Roman Bold	**اشكرك**	0.14	0.01	**	0.00	0.94		0.05	0.29	
Traditional Arabic	اشكرك	0.18	0.001	***	0.04	0.39		0.12	0.02	*
Traditional Arabic Bold	**اشكرك**	0.22	0.001	***	0.07	0.15		0.15	0.001	***
Urdu Typesetting	اشكرك	0.07	0.18		0.07	0.18		0.02	0.65	
Urdu Typesetting Bold	**اشكرك**	0.05	0.29		0.04	0.39		0.01	0.78	

Note. * $p < .05$, ** $p < .01$, *** $p < .001$. The phonological distances were measured by cosine distance.

Table 3B. Grapho-phonemic systematicity of 39 Hebrew fonts.

		Pixel count			Perimetric complexity			Hausdorff distance		
Font	Example	r	p		r	p		r	p	
Aharoni	שלום	0.20	0.001	***	0.10	0.07		0.14	0.01	**
Arial	שלום	0.18	0.001	***	0.14	0.01	**	0.12	0.03	*
Arial Bold	שלום	0.19	0.001	***	0.11	0.06		0.10	0.06	
Arial Bold Italic	שלום	0.19	0.001	***	0.15	0.01	**	0.11	0.04	*
Arial Italic	שלום	0.17	0.001	***	0.20	0.001	***	0.12	0.03	*
Calibri	שלום	0.19	0.001	***	0.14	0.01	**	0.14	0.01	**
Calibri Bold	שלום	0.18	0.001	***	0.16	0.001	***	0.14	0.01	**
Calibri Italic	שלום	0.19	0.001	***	0.20	0.001	***	0.13	0.02	*
Calibri Light	שלום	0.18	0.001	***	0.16	0.001	***	0.13	0.02	*
Calibri Light Italic	שלום	0.19	0.001	***	0.20	0.001	***	0.13	0.02	*
Calibri Bold Italic	שלום	0.19	0.001	***	0.20	0.001	***	0.14	0.01	**
Courier New	שלום	0.22	0.001	***	0.17	0.001	***	0.17	0.001	***
Courier New Bold	שלום	0.21	0.001	***	0.15	0.01	**	0.19	0.001	***
Courier New Bold Italic	שלום	0.20	0.001	***	0.22	0.001	***	0.19	0.001	***
Courier New Italic	שלום	0.19	0.001	***	0.22	0.001	***	0.18	0.001	***
David	שלום	0.18	0.001	***	0.20	0.001	***	0.11	0.05	*
David Bold	שלום	**0.19**	0.001	***	0.15	0.01	**	0.10	0.07	
Franklin Gothic Book	שלום	0.19	0.001	***	0.19	0.001	***	0.13	0.02	*
Gisha	שלום	0.15	0.01	**	0.12	0.03	*	0.14	0.01	**
Gishabd	שלום	0.19	0.001	***	0.12	0.03	*	0.14	0.01	**
Levenim MT	שלום	0.12	0.04	*	0.09	0.12		0.13	0.02	*
Levenim MT Bold	שלום	0.15	0.01	**	0.09	0.10		0.13	0.02	*
Lucida Sans Unicode	שלום	0.14	0.01	**	0.13	0.02	*	0.13	0.02	*
Microsoft Sans Serif	שלום	0.14	0.01	**	0.11	0.05	*	0.11	0.06	
Miriam	שלום	0.16	0.001	***	0.13	0.02	*	0.10	0.09	
Miriam Fixed	שלום	0.15	0.01	**	0.13	0.02	*	0.14	0.01	**
Narkisim	שלום	0.19	0.001	***	0.15	0.01	**	0.11	0.05	*
Rod	שלום	0.20	0.001	***	0.16	0.001	***	0.12	0.03	*
Segoe UI	שלום	0.11	0.04	*	0.07	0.23		0.12	0.03	*
Segoe UI Bold	שלום	0.13	0.02	*	0.06	0.32		0.11	0.05	*
Segoe UI Light	שלום	0.09	0.09		0.09	0.10		0.11	0.05	*
Segoe UI Semilight	שלום	0.12	0.03	*	0.09	0.12		0.11	0.04	*
Segoe UI Semibold	שלום	0.13	0.02	*	0.13	0.02	*	0.11	0.05	*
Tahoma	שלום	0.12	0.03	*	0.07	0.20		0.12	0.04	*

Tahoma Bold	**שלום**	0.13	0.02	*	0.09	0.13		0.11	0.04	*
Times New Roman	שלום	0.14	0.01	**	0.18	0.001	***	0.13	0.02	*
Times New Roman Bold	**שלום**	0.12	0.03	*	0.17	0.001	***	0.12	0.03	*
Times New Roman Bold Italic	*שלום*	0.15	0.01	**	0.15	0.01	**	0.14	0.01	**
Times New Roman Italic	*שלום*	0.12	0.03	*	0.20	0.001	***	0.13	0.02	*

Note. * p < .05, ** p < .01, *** p < .001. The phonological distances were measured by feature edit distance.

Table 3C. Grapho-phonemic systematicity of the 24 most frequent English fonts.

	Pixel count						Perimetric complexity						Hausdorff distance					
	Upper			Lower			Upper			Lower			Upper			Lower		
Font	*r*	*p*		*r*	*p*		*r*	*p*		*r*	*p*		*r*	*p*		*r*	*p*	
Arial	0.20	0.001	***	0.12	0.04	*	0.10	0.08		0.07	0.27		0.01	0.84		0.09	0.15	
Arial Black	0.20	0.001	***	0.16	0.01	***	-0.04	0.55		-0.03	0.65		0.06	0.33		0.14	0.02	*
Calibri	0.26	0.001	***	0.11	0.06		0.16	0.01	*	0.01	0.82		0.09	0.12		0.09	0.16	
Cambria	0.18	0.001	***	0.14	0.02	*	0.20	0.001	***	0.10	0.10		0.12	0.04	*	0.13	0.03	*
Candara	0.24	0.001	***	0.09	0.13		0.22	0.001	***	0.06	0.33		0.08	0.20		0.11	0.07	
Comic Sans MS	0.26	0.001	***	0.11	0.07		0.11	0.07		0.07	0.27		0.01	0.84		-0.02	0.76	
Consolas	0.09	0.14		0.001	0.94		0.19	0.001	***	0.10	0.11		-0.01	0.93		-0.01	0.84	
Constantia	0.29	0.001	***	0.11	0.07		0.20	0.001	***	0.13	0.04	*	0.10	0.09		0.08	0.19	
Corbel	0.25	0.001	***	0.12	0.05	*	0.22	0.001	***	0.15	0.01	**	0.05	0.42		0.05	0.43	
Courier New	0.11	0.06		-0.01	0.86		0.23	0.001	***	0.03	0.59		0.001	1.00		-0.05	0.44	
Franklin Gothic Medium	0.21	0.001	***	0.20	0.001	***	0.26	0.001	***	-0.01	0.92		0.06	0.31		0.15	0.01	**
Gabriola	0.21	0.001	***	0.06	0.32		0.24	0.001	***	-0.03	0.61		0.12	0.04	*	-0.03	0.66	
Georgia	0.23	0.001	***	0.17	0.01	**	0.21	0.001	***	0.09	0.13		0.10	0.09		0.05	0.39	
Impact	0.20	0.001	***	0.18	0.001	***	0.10	0.09		0.12	0.04	*	0.15	0.01	**	0.10	0.08	
Lucida Console	0.14	0.02	*	0.01	0.85		0.10	0.08		0.17	0.01	**	-0.02	0.73		0.01	0.83	
Microsoft Sans Serif	0.18	0.001	***	0.15	0.01	**	0.14	0.02	*	0.06	0.36		0.001	0.97		0.06	0.32	
Palatino Linotype	0.20	0.001	***	0.14	0.02	*	0.25	0.001	***	0.04	0.50		0.09	0.12		0.06	0.32	
Segoe Print	0.12	0.05	*	0.18	0.001	***	0.18	0.001	***	0.17	0.001	***	0.001	0.98		0.12	0.05	*
Segoe Script	0.06	0.34		0.03	0.63		0.18	0.001	***	-0.06	0.34		0.03	0.61		0.03	0.59	
Segoe UI Symbol	0.24	0.001	***	0.12	0.05	*	0.13	0.03	*	0.14	0.02	*	0.08	0.21		0.03	0.59	
Tahoma	0.18	0.001	***	0.18	0.001	***	0.08	0.17		0.12	0.05	*	0.06	0.36		0.12	0.04	*
Times New Roman	0.14	0.02	*	0.15	0.01	**	0.20	0.001	***	0.07	0.25		0.05	0.42		0.08	0.21	
Trebuchet MS	0.18	0.001	***	0.14	0.02	*	0.2	0	***	0.1	0.09		0.05	0.43		0.08	0.2	
Verdana	0.17	0.001	***	0.16	0.01	**	0.03	0.65		0.11	0.08		0.00	0.94		0.14	0.02	*

Note. * p < .05, ** p < .01, *** p < .001. The phonological distances were measured by feature edit distance.

Table 3C shows consistent systematicity for the uppercase letters and less consistent systematicity for the lowercase letters.

Discussion

We have presented a proof of concept—that quantifying the comprehensive relation between letter-shapes and their canonical pronunciations across a whole alphabet is a tractable and instructive task. Crucially, we have shown that characters or letters that look similar tend to have similar canonical pronunciations. We found examples of strong grapho-phonemic systematicity in three orthographies (Arabic, Hebrew and English) that have evolved culturally over a long period and in three orthographies (Korean, Pitman's shorthand and the Shavian alphabet) that were intentionally developed over a short period to be systematic. The latter represent a point of comparison for the former.

Our results bear on studies that have examined the cultural evolution of writing systems, in which the researchers have abstracted away from the comprehensive relationship of all the characters with their canonical phonology across the whole of the particular writing system. We suggest that this fuller context should be considered in studies of dimensions such as graph inventory size and overall type of writing system (Chang, Plaut & Perfetti, 2016; Chang, Chen, and Perfetti, 2018; Changizi & Shimojo, 2005), phylogeny (Milton & Morin, 2021), information density (cf. Jaeger, 2010), homogeneity between characters (Treiman & Kessler, 2011), graphic complexity (Pelli et al., 2006; Tamariz & Kirby, 2015), and reading direction (Fischer, 2011). In the case of grapho-phonemic orthographies, studying isolated individual letters does not tell the full story.

Our results for the six orthographies all show grapho-phonemic systematicity. However, they also suggest that applications of the approach to other orthographies and languages may reveal quantitative and perhaps qualitative differences in the nature of such systematicity.

We used a suite of three measures of visual distance. Our initial motivation was that researchers in language systematicity have argued that systematicity shown by multiple metrics is more compelling as it tends to compensate for the idiosyncrasies of particular methods. This reasoning is particularly important when a new phenomenon is being explored.

Two things are clear from Tables 1-3(C). First, the three visual measures did not reveal equivalent degrees of systematicity when correlated against the same measure of phonological distance. Second, the three measures behaved differently with respect to the two groups of writing systems.

<u>With respect to the first point,</u> above, Korean, Pitman's shorthand, and the Shavian Alphabet—all consciously created as systems— exhibited significant systematicity. Their authors consciously generated general visual symbols for speech sounds, as single projects over a short period of time. Similar phonemes were intentionally designed to look similar. For example, Korean aspirated phonemes have one additional stroke compared with the lenis phonemes (e.g. ㅋ /k/ -- ㄱ /g/). Pitman's shorthand distinguishes English voiced and voiceless phonemes using line thickness; the Shavian alphabet makes the same distinction using rotation. We term this type of regularity 'symbol-based grapho-phonemic systematicity'. Table 1 shows that this symbol-based (or 'featural') systematicity was more fully revealed by the Hausdorff visual distance metric. The detailed topographic nature of this method—capturing between-letter similarities in line length, orientation and position—successfully reveals the systematic letter variation that is

correlated with phonological distance. Pixel count and perimetric complexity cannot capture this level of detail.

Pixel count and perimetric complexity both reflect the number of pixels in a letter, but they differ in their capacity to reflect shape. Pixel count is identical for a 20-pixel black blob and a 20-pixel thin black line, but the perimetric complexity values are very different. Thus, perimetric complexity is sensitive to line thickness. Pitman's shorthand does in fact use thickness of line as a systematic feature (see Appendix1). In Table 1, we can see perimetric complexity eliciting a higher correlation than pixel count for Pitman's shorthand.

Of the three writing systems in Table 1, only Korean presented us with a choice of font. Table 2 gives examples of the level of systematicity for the individual Korean fonts. In Table 1 the mean Korean systematicity was calculated using the mean of the different visual distances between particular pairs of letters over all of the fonts. Using all the readily available fonts ensures the robustness of the results. Table 2 shows that, as in all our results for the real languages, the overall pattern is graded and a minority of fonts go in the opposite direction.

Second, as noted above, the three measures behaved differently with respect to the two groups of writing systems. We studied systematicity in the symbol-based writing systems so as to provide a baseline for exploring systematicity in our three orthographies that evolved piecemeal over centuries—Arabic, Hebrew and English. We knew that systematicity had been purposely constructed in the first group. Could comprehensively quantified systematicity in our three 'naturally occurring' writing systems reach comparable levels? Across the three orthographies, it was captured best by simple pixel count and least well by Hausdorff distance. Again, we see a graded pattern of significant corelations, both within and between languages. Languages that have evolved piecemeal over an extended cultural period can and do include symbolic or ‘featural’ regularities (e.g. b /b/ and p /p/ in English; ח /x/ -- ה /h/ in Hebrew) but instances tend to be fewer and less consistent, compared with the intentionally constructed writing systems.

English uppercases exhibited more systematicity than lowercases. Lowercase followed uppercase historically (Harris, 2003). It also varies more across different fonts; compare ‘a’ and ‘A’ in Times Roman with ‘a’ and ‘A’ in Comic Sans. This factor would affect systematicity if averaged over many fonts. However, uppercases still tend to be more systematic for individual fonts (Table 3C). We conclude that the lowercase adaptations to cursive writing have obscured some of the systematicity found in uppercase.

Because pixel count is the least sensitive to structure, we term the systematicity most in evidence in Arabic, Hebrew and English ‘effort-based grapho-phonemic systematicity’. Some speech sounds are ‘more elaborate’, in that they require more articulatory effort to produce and more processing effort to perceive because of intrinsic and relative aspects of manner and/or place of articulation. Phonemes that are more elaborate to produce tend to be represented by letters that are also more elaborate to produce, where ‘more elaborate’ means more pixels and thus more *effort* to reproduce the letter or character. In this aspect of systematicity, there is only one, continuous ‘feature’—the degree of effort involved.

Phonological elaborateness tends to be proportional to number of pixels. Because we have defined phonological distance in terms of the featural differences between two phonemes, less effortful phonemes tend to be closer together in phonological space. Analogously, letters that are simpler and less effortful to write tend to be closer together in any graphological space that reflects the amount of writing necessary (e.g. number of pixels). From this effort-based perspective in processing phonology and in creating letters, a correlation between the two naturally emerges, resulting in an ‘effort-based systematicity’ distributed across the alphabet and the corresponding canonical pronunciations—this is the nature of the systematicity we have reported for Arabic, Hebrew and English.

Our metrics have been sufficient to demonstrate two tendencies within systematicity. One, symbol-based, exhibits the compositionality that we see in other subdomains of language use, such as morphology and syntax. The other, effort-based, is more basic but is also seen in other aspects of language use, as in Zipf’s (1949) emphasis on effort in language. How important might these two tendencies within systematicity be for using writing systems? Any use of an orthography must be learned. We suggest that the structure of grapho-phonemic systematicity in a writing-system points to an optimum order of letters by which beginning readers might be introduced to the alphabetic principle in the phonics approach. (See Jee, 2021, for the ranking of characters inside an orthography in terms of how much they individually contribute to overall systematicity.)

The order of the early learning of canonical pronunciations of letters may not be entirely superseded by subsequent experience (cf. Brown & Watson, 1987; Pelli, Farell, & Moore, 2003), raising the question of whether grapho-phonemic systematicity plays a role in adult skilled reading. There is a vast literature on the automatic uptake of phonology in reading, with key questions concerning the processing of parafoveal words and letters (cf. Schotter, Angele, & Rayner, 2012; Tiffin-Richards, & Schroeder, 2015; Tsai, Lee, Tzeng, Hung, & Yen, 2004; Yan, Wang, Song, & Kliegl, 2019). We suggest that it is in this fragile aspect of skilled reading that what may be a small effect of grapho-phonemic systematicity is most likely to be visible.

There is ongoing research to assess claims concerning optimal fonts for dyslexic readers (e.g. Bachmann & Mengheri, 2018; Kuster, van Weerdenburg, Gompel, & Bosman, 2018;

Marinus et al., 2016; Rello & Baeza-Yates, 2013). We propose that quantified systematicity is one objective criterion for investigating particular fonts. Higher systematicity should facilitate letter and word processing, given that suboptimal phonological processing is an influential explanation for the reading problems of a proportion of dyslexics (cf. Ramus et al., 2003), as is a deficit in the automatization of verbal responses to visual stimuli (cf. Denckla & Rudel, 1976).

Behind these issues of reading behaviour is a more fundamental question of the cognitive neuropsychological instantiation of systematicity. Based on the characteristic processing propensities of the left and right cerebral hemispheres (cf. McGilchrist's 2021 extensive review), we suggest that grapho-phonemic systematicity has different implications for the two hemispheres. The right hemisphere's use of context is well-suited to recognising systematicity in the first place. The right hemisphere's involvement with 'authenticity' perhaps also underwrites readers' experience that there is something '/b/-like' about the letter "b". In contrast, the left hemisphere's strength in the automatic processing of representations lends itself to the fast, predictive matching of graphemes with phonemes.

We have reported computations performed over quantified images and theoretical analyses of speech sounds. We have not reported laboratory-based behavioural data. However, our data *are* behavioural in a different sense: the behaviours of speaking, listening, and making marks; the behaviours of the authors of new writing systems; the behaviours of the transcribers, writers and educators who learned a writing system and then passed it on to the next generation. Our data speak profoundly to questions of cognition and action.

The psychological reality of our metrics is an empirical question. However, it is not a critical one. What *is* critical is that we have quantified the existence of significant systematicity. Any systematicity we can reveal using very simple metrics is certainly also accessible to the human brain. Other metrics (perhaps particularly ones based on observed effort) will almost certainly be able to demonstrate even greater systematicity.

However, it is initially misleading to consider a factor such as the phonological complexity of a linguistic entity in isolation from other factors, notably the *frequency* of the entity and the *complexity of its context*. In any real-world processing, all these factors are present in one totality. As we will see below, the intercorrelation of such factors suggests an understanding of the nature of other cases of systematicity in language and eventually provides us with the most parsimonious understanding of all such systematicity.

As a first example of such an intercorrelation, the greater frequency of occurrence of a linguistic entity can be seen as ‘more opportunities to occur in a greater range of contexts’. Thus the correlated dimensions of frequency and contextual variability can be *proxies* for each other. Word frequency effects in language processing have been conventionally understood in terms of Hebbian learning (Hebb, 1949); words that are encountered more often become easier to process because the relevant neural pathways are facilitated by frequent use. However, McDonald and Shillcock (2001) show that a measure of contextual variability, “Contextual Distinctiveness”, provides a better account of the behavioural data than does frequency.

As a second example of intercorrelated dimensions, consider a phoneme’s ‘phonotactic range’, which is the variability of its context—the number of different phonemes that can appear adjacent to it in speech. Phonotactic range and phoneme frequency both predict infants’ order of phonological acquisition in English and the former is actually the better predictor (Shillcock & Westermann, 1996): segments occurring next to a greater variety of other segments in speech are deployed earlier in infancy.

Segment complexity is also part of this picture: a segment with fewer articulatory constraints (i.e. features) is able to occur next to more segment types, simply because of the mechanics of speech production. Thus, segment complexity tends to be inversely correlated with variability of context and with segment frequency. Segment frequency, intrinsic featural complexity and variability of context are all highly intercorrelated.

These considerations of the nature of systematicity in language processing lead to our final question—does our effort-based interpretation of grapho-phonemic systematicity have implications for the influential research concerning *phono-semantic* systematicity, in which *whole words* that sound similar tend to have similar meanings? Recall that this effect is at a distributed, 'sub-symbolic' level, not the symbolic level of morphemes and phonoaesthetic effects. Elsewhere, we have demonstrated phono-semantic systematicity in Korean (Jee, Tamariz & Shillcock, 2022). Crucially, we have shown that the effect persists even when the sample words are partitioned in terms of etymological origin (Korean or Sino-Korean), syntactic category, homonyms, onomatopoeia, loanwords, and even in terms of syllabic constituents (onset, vowel, coda, and rhyme). The systematicity is pervasive. It only falls below statistical significance in the stratum of the lowest frequency words. We claim here that the role of *cognitive effort* unites that study and the current one: it is adaptive for frequent types to be less effortful, whether they are letters, words, pronunciations or meanings.

Specifically, more frequent words are shorter and with a quantifiably less complex phonology. More frequent words occur in a greater variety of contexts (cf. our discussion of phonotactic range, above) and acquire more densely populated, self-similar LSA (Latent Semantic Analysis; Landauer & Dumais, 1997) context vectors standing for meaning/usage. A form-meaning systematicity *necessarily* emerges from this correlation; as words gets longer and more complex, their context vectors become sparser and more differentiated, with corresponding increases in phonological and ‘semantic’ distances between words. The overall form-meaning relation in the monomorphemic words of any naturally occurring language is thus not arbitrary—words that sound similar tend to have similar meanings—but it is inescapable.

We can bring this analysis of phono-semantic systematicity into line with our effort-based theory of the relation between letter sounds and shapes. In a *complex system* everything interacts with everything else. Processing is *less effortful* for frequent types, which are intrinsically simpler and are supported by many default relations between many entities. Processing is *more effortful*, for infrequent types, which are intrinsically more complex and require many default relations between many entities to be inhibited to allow a sparse, idiosyncratic constellation of relations to emerge. We propose that this effort-based account is the underlying explanation of phono-semantic systematicity.

The strength of such an explanation is that it is expressed in terms of the lowest level processing instantiated across the widest cognitive substrate. The lowest level processing component is ‘effort required’ and the widest cognitive substrate involves large areas of the brain plus those parts of the body in which cognition is *embodied* (i.e. the organs responsible for speech, the hand for writing).

In conclusion, in our research on grapho-phonemic (and phono-semantic systematicity) we have returned to Zipf’s (1949) emphasis on effort in language, but with new methodologies, new data, new phenomena, and with our modelling and theorizing enriched by research on complex systems.

REFERENCES

Algeo, J., & Butcher, C. A. (2013). *The origins and development of the English language*. Cengage Learning.

Bachmann, C., & Mengheri, L. (2018). Dyslexia and fonts: Is a specific font useful?. *Brain Sciences*, *8*(5), 89.

Berndt, R. S., Reggia, J. A., & Mitchum, C. C. (1987). Empirically derived probabilities for grapheme-to-phoneme correspondences in English. *Behavior Research Methods, Instruments, & Computers*, *19*(1), 1-9.

Blank, D. S., Meeden, L. A., & Marshall, J. B. (1992). Exploring the symbolic/subsymbolic continuum: A case study of RAAM. In J. Dinsmore (Ed.), *The symbolic and connectionist paradigms: Closing the gap* (pp. 113–148). Lawrence Erlbaum Associates, Inc.

Blasi, D. E., Wichmann, S., Hammarström, H., Stadler, P. F., & Christiansen, M. H. (2016). Sound–meaning association biases evidenced across thousands of languages. *Proceedings of the National Academy of Sciences*, *113*(39), 10818-10823.

Brown, G. D., & Watson, F. L. (1987). First in, first out: Word learning age and spoken word frequency as predictors of word familiarity and word naming latency. *Memory & cognition*, *15*(3), 208-216.

Castles, A., Rastle, K., & Nation, K. (2018). Ending the reading wars: Reading acquisition from novice to expert. *Psychological Science in the Public Interest*, *19*(1), 5-51.

Chang, W. D., Kim, H. Y., Cha, E. Y., & Kim, D. H. (2005). The Recognition of Grapheme'ㅁ','ㅇ'Using Neighbor Angle Histogram and Modified Hausdorff Distance. *Journal of Korea Multimedia Society*, *8*(2), 181-191.

Chang, L. Y., Chen, Y. C., & Perfetti, C. A. (2018). GraphCom: A multidimensional measure of graphic complexity applied to 131 written languages. *Behavior research methods*, *50*(1), 427-449.

Chang, L. Y., Plaut, D. C., & Perfetti, C. A. (2016). Visual complexity in orthographic learning: Modeling learning across writing system variations. *Scientific Studies of Reading*, *20*(1), 64-85.

Changizi, M. A., & Shimojo, S. (2005). Character complexity and redundancy in writing systems over human history. *Proceedings of the Royal Society B: Biological Sciences*, *272*(1560), 267-275.

Dautriche, I., Mahowald, K., Gibson, E., & Piantadosi, S. T. (2017). Wordform similarity increases with semantic similarity: An analysis of 100 languages. *Cognitive Science*, *41*(8), 2149-2169.

Denckla, M. B., & Rudel, R. G. (1976). Rapid ‘automatized’naming (RAN): Dyslexia differentiated from other learning disabilities. *Neuropsychologia*, *14*(4), 471-479.

Dingemanse, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., & Monaghan, P. (2015). Arbitrariness, iconicity, and systematicity in language. *Trends in Cognitive Sciences*, *19*(10), 603-615.

Fischer, J. P. (2011). Mirror writing of digits and (capital) letters in the typically developing child. *Cortex*, *47*(6), 759-762.

Fitneva, S. A., Christiansen, M. H., & Monaghan, P. (2009). From sound to syntax: Phonological constraints on children's lexical categorization of new words. *Journal of Child Language*, *36*(5), 967-997.

Frost, R. (1995). Phonological computation and missing vowels: Mapping lexical involvement in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(2), 398.

Geyer, L. H. (1977). Recognition and confusion of the lowercase alphabet. *Perception & Psychophysics*, *22*(5), 487-490.

Gibson, E. P., Gibson, J. J., Pick, A. D., & Osser, H. (1962). A developmental study of the discrimination of letter-like forms. *Journal of Comparative and Physiological Psychology*, *55*(6), 897–906.

Gontijo, P. F., Gontijo, I., & Shillcock, R. (2003). Grapheme—phoneme probabilities in British English. *Behavior Research Methods, Instruments, & Computers*, *35*(1), 136-157.

Harm, M. W., & Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: insights from connectionist models. *Psychological Review*, *106*(3), 491-528.

Harris, D. (2003). *The Calligrapher's Bible*. Hauppauge, NY: Barron's. ISBN 0-7641-5615-2.

Hebb, D. O. (1949). *The organisation of behaviour: a neuropsychological theory*. New York: Science Editions.

Hinton, L., Nichols, J., & Ohala, J. J. (Eds.). (2006). *Sound symbolism*. Cambridge University Press.

Huttenlocher, D. P., Klanderman, G. A., & Rucklidge, W. J. (1993). Comparing images using the Hausdorff distance. *IEEE Transactions on pattern analysis and machine intelligence*, *15*(9), 850-863.

Jaeger, T. F. (2010). Redundancy and reduction: Speakers manage syntactic information density. *Cognitive psychology*, *61*(1), 23-62.

Jee, H., Tamariz, M. & Shillcock, R. (2020). Quantifying Sound-Graphic Systematicity: Application to Multiple Phonographic orthographies. *Proceedings Grapholinguistics and Its Application*, *4a*, 123–143.

Jee, H. (2021). *Linguistic systematicity between phonology, semantics and orthography in English and Korean*. Unpublished PhD dissertation, University of Edinburgh.

Jee, H., Tamariz, M. & Shillcock, R. (2022). Exploring meaning-sound systematicity in Korean. *Journal of East Asian Linguistics, 30*(1)

Kelly, M. H. (1992). Using sound to solve syntactic problems: The role of phonology in grammatical category assignments. *Psychological Review*, *99*(2), 349-364.

Kriegeskorte, N., Mur, M., & Bandettini, P. A. (2008). Representational similarity analysis-connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, *2*, 4.

Kuster, S. M., van Weerdenburg, M., Gompel, M., & Bosman, A. M. (2018). Dyslexie font does not benefit reading in children with or without dyslexia. *Annals of Dyslexia*, *68*(1), 25-42.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, *104*(2), 211-240.

Lloyd, S., Wernham, S., Jolly, C., & Stephen, L. (1998). *The phonics handbook.* Chigwell: Jolly Learning.

Marinus, E., Mostard, M., Segers, E., Schubert, T. M., Madelaine, A., & Wheldall, K. (2016). A special font for people with dyslexia: Does it work and, if so, why?. *Dyslexia*, *22*(3), 233-244.

McDonald, S. A., & Shillcock, R. C. (2001). Rethinking the word frequency effect: The neglected role of distributional information in lexical processing. *Language and Speech*, *44*(3), 295-322.

McGilchrist, I. (2021). *The Matter With Things: Our Brains, Our Delusions and the Unmaking of the World*. Perspective Press, London.

Monaghan, P., Shillcock, R. C., Christiansen, M. H., & Kirby, S. (2014). How arbitrary is language?. *Phil. Trans. R. Soc. B*, *369*(1651), 20130299

Morgan, J. L., & Demuth, K. (1996). Signal to syntax: An overview. *Signal to syntax: Bootstrapping from speech to grammar in early acquisition*, (Erlbaum, Hillsdale, NJ). pp. 1-22.

Miton, H., & Morin, O. (2021). Graphic complexity in writing systems. *Cognition*, *214*, 104771.

Pelli, D. G., Burns, C. W., Farell, B., & Moore-Page, D. C. (2006). Feature detection and letter identification. *Vision research, 46*(28), 4646-4674.

Pelli, D. G., Farell, B., & Moore, D. C. (2003). The remarkable inefficiency of word recognition. *Nature*, *423*(6941), 752-756.

Pitman, I. (1845). *A Manual of Phonography; or, Writing by Sound* (7 ed.). London: S. Bagster.

Ramus, F., Rosen, S., Dakin, S. C., Day, B. L., Castellote, J. M., White, S., & Frith, U. (2003). Theories of developmental dyslexia: insights from a multiple case study of dyslexic adults. *Brain*, *126*(4), 841-865.

Rello, L., & Baeza-Yates, R. (2013, October). Good fonts for dyslexia. In *Proceedings of the 15th international ACM SIGACCESS conference on computers and accessibility* (pp. 1-8).

Robinson, A. (1995). *The story of writing: Alphabets, hieroglyphs & pictograms*. New York: Thames and Hudson.

Sampson, G. (1985). *Writing systems*. London, U K: Hutchinson.

Sanders, N. C., & Chin, S. B. (2009). Phonological distance measures. *Journal of quantitative linguistics*, *16*(1), 96-114.

Schotter, E. R., Angele, B., & Rayner, K. (2012). Parafoveal processing in reading. *Attention, Perception, & Psychophysics*, *74*(1), 5-35.

Seymour, P. H., Aro, M., Erskine, J. M., & collaboration with COST Action A8 network. (2003). Foundation literacy acquisition in European orthographies. *British Journal of psychology*, *94*(2), 143-174.

Shi, R., Morgan, J. L., & Allopenna, P. (1998). Phonological and acoustic bases for earliest grammatical category assignment: A cross-linguistic perspective. *Journal of Child Language*, *25*(1), 169-201.

Shillcock, R., Kirby, S., McDonald, S., & Brew, C. (2001). Filled pauses and their status in the mental lexicon. In *ISCA Tutorial and Research Workshop (ITRW) on Disfluency in Spontaneous Speech*, 53–56.

Shillcock, R., & Westermann, G. (1996). The role of phonotactic range in the order of acquisition of English consonants. *Proceedings of the ICPLA*.

Simons, F. (2011). Proto-Sinaitic–Progenitor of the Alphabet. *Rosetta*, *9*, 16-40.

Smolensky, P., & Prince, A. (1993). Optimality Theory: Constraint interaction in generative grammar. *Optimality Theory in phonology*, 3.

Song, H. J., & Wiese, R. (2010). Resistance to complexity interacting with visual shape—German and Korean orthography. *Writing Systems Research*, *2*(2), 87-103.

Tamariz, M. (2008). Exploring systematicity between phonological and context-cooccurrence representations of the mental lexicon. *The Mental Lexicon*, *3*(2), 259-278.

Tamariz, M., & Kirby, S. (2015). Culture: copying, compression, and conventionality. *Cognitive science*, *39*(1), 171-183.

Tiffin-Richards, S. P., & Schroeder, S. (2015). Children's and adults' parafoveal processes in German: Phonological and orthographic effects. *Journal of Cognitive Psychology*, *27*(5), 531-548.

Treiman, R., & Kessler, B. (2011). Similarities among the shapes of writing and their effects on learning. *Written Language & Literacy*, *14*(1), 39-57.

Tsai, J. L., Lee, C. Y., Tzeng, O. J., Hung, D. L., & Yen, N. S. (2004). Use of phonological codes for Chinese characters: Evidence from processing of parafoveal preview when reading sentences. *Brain and Language*, *91*(2), 235-244.

Yan, M., Wang, A., Song, H., & Kliegl, R. (2019). Parafoveal processing of phonology and semantics during the reading of Korean sentences. *Cognition*, *193*, 104009.

Zipf, G. K. (1949/2016). *Human behavior and the principle of least effort: An introduction to human ecology*. Ravenio Books.

Appendix 1.

Pitman’s shorthand

letter						
name	pee	bee	tee	dee	chay	jay
phoneme	/p/	/b/	/t/	/d/	/tʃ/	/dʒ/
letter						
name	kay	gay	ell	ar	way	yay
phoneme	/k/	/g/	/l/	/r/	/w/	/j/
						
name	em	en	eff	vee	ith	thee
phoneme	/m/	/n/	/f/	/v/	/θ/	/ð/
						
name	ing	hay	ess	zee	ish	zhee
phoneme	/h/	/ŋ/	/s/	/z/	/ʃ/	/ʒ/

Shavian alphabet

letter	𐑐	𐑚	𐑑	𐑛	𐑗	𐑡
name	peep	bib	tot	dead	church	judge
phoneme	/p/	/b/	/t/	/d/	/tʃ/	/dʒ/
letter	𐑒	𐑜	𐑤	𐑮	𐑢	𐑘
name	kick	gag	loll	roar	woe	yea
phoneme	/k/	/g/	/l/	/r/	/w/	/j/
letter	𐑥	𐑯	𐑓	𐑝	𐑔	𐑞
name	mime	nun	fee	vow	thigh	they
phoneme	/m/	/n/	/f/	/v/	/θ/	/ð/
letter	𐑣	𐑙	𐑕	𐑟	𐑖	𐑠
name	ha-ha	hung	so	zoo	sure	measure
phoneme	/h/	/ŋ/	/s/	/z/	/ʃ/	/ʒ/