

Character complexity and redundancy in writing systems over human history

Mark A. Changizi^{1*} and Shinsuke Shimojo^{2,3}

¹Sloan-Swartz Center for Theoretical Neurobiology, and ²Division of Biology, Computation and Neural Systems, MC 139-74, Caltech, Pasadena, CA 91125, USA ³NTT Communication Science Laboratory, Atsugi, Kanagawa, Japan

A writing system is a visual notation system wherein a repertoire of marks, or strokes, is used to build a repertoire of characters. Are there any commonalities across writing systems concerning the rules governing how strokes combine into characters; commonalities that might help us identify selection pressures on the development of written language? In an effort to answer this question we examined how strokes combine to make characters in more than 100 writing systems over human history, ranging from about 10 to 200 characters, and including numerals, abjads, abugidas, alphabets and syllabaries from five major taxa: Ancient Near-Eastern, European, Middle Eastern, South Asian, Southeast Asian. We discovered underlying similarities in two fundamental respects.

- (i) The number of strokes per characters is approximately three, independent of the number of characters in the writing system; numeral systems are the exception, having on average only two strokes per character.
- (ii) Characters are *ca*. 50% redundant, independent of writing system size; intuitively, this means that a character's identity can be determined even when half of its strokes are removed.

Because writing systems are under selective pressure to have characters that are easy for the visual system to recognize and for the motor system to write, these fundamental commonalities may be a fingerprint of mechanisms underlying the visuo–motor system.

Keywords: writing; reading; redundancy; complexity; letter perception; character recognition

1. INTRODUCTION

Writing systems (such as alphabets) are visual notation systems wherein a repertoire of marks, or strokes, is used to construct a set of characters. Because writing systems are under selective pressure to be easy to read and write, we reasoned that by identifying commonalities underlying how strokes combine into characters in writing systems over human history, we would, in effect, be identifying fundamental properties of the human visuo-motor system. Here, we are interested in measuring two specific fundamental properties of writing systems: character length, which is the average number of strokes per character, and redundancy, which measures how efficiently characters are built out of strokes. Our main result will be that, after examining more than 100 writing systems over human history (table 1), ranging from about 10 to 200 characters, we determined that writing systems have average lengths of approximately three strokes per character and redundancies of ca. 50%, and these values do not vary much as a function of writing system size. In $\S 4$ we will speculate on what this might tell us about the human visuo-motor system.

2. RESULTS

Suppose a writing system possesses B stroke types, and average character length L. We can model the number of characters, C, by an equation of the general form,

 $C = \sigma B^{\beta L}$, where $\sigma, \beta \leq 1$ are positive constants. The proportionality constant, σ , captures the fact that some fraction of the possible stroke combinations may not be allowed. The exponent βL is called the *combinatorial degree*, d, and measures how combinatorially strokes are used to build characters: a minimum value of d = 1 would mean that strokes are not used combinatorially at all (i.e. because doubling the number of stroke types would only double the number of characters), and greater values (up to a maximum of L) mean that strokes are used more combinatorially. From the length, L, and combinatorial degree, d, we may compute the redundancy, R = 1 - (d/L), or $R = 1 - \beta$: a redundancy of zero means that all L potential degrees of freedom in building characters are used, and as the redundancy nears 1 the combinatorial degree falls lower and lower below L. For example, suppose that 0s and 1s can be strung together to make sequences of length 4, but that 0s must always be placed next to a 0, and 1s next to a 1. Although there are $2^4 = 16$ binary sequences of length 4, there are only C = 4 sequences satisfying this constraint, namely 0000, 0011, 1100 and 1111. Here, $\sigma = 1$ and $\beta = 1/2$, so that $C = 4 = 1 \times 2^{(1/2)4} = \sigma B^{\beta L}$. Intuitively, the constraints put on these sequences reduces the number of degrees of freedom from 4 down to 2, and thus the sequence length is twice as long as it ideally would have to be, corresponding to a redundancy of 1/2. The three properties-combinatorial degree d, length L, and redundancy *R*—are related such that any two are independent,

^{*}Author for correspondence (changizi@changizi.com).

Table 1. Information for the writing systems used in the analysis.

so as to cover all major phylogenetically distinct writing system classes for numerals, abjads, abugidas (including vowel diacritics), alphabets and syllabaries, as shown in figure 1. Invented writing systems were included if they have been successfully used by some community; writing systems invented for fun or for fictional purposes were not included. We did not include logotiple variants existed, we chose the first shown. Lower-case characters were used unless otherwise noted; also, character variants that occur at the start or end of a word were not used. In alphabets, the alphabet has some syllabic features, and 'syllabic' has been placed in parentheses to denote this. Dates are often only approximate, and are recorded such that negative age character length, which is the average number of strokes per character in the writing system (see figure 2*a*). The sixth column gives the stroke type repertoire size (see figure 3*a*), and the seventh column the writing system size (i.e. the number of characters). The last column provides the total number of edges in the stroke-type network for that writing system (see text and (Character information was sampled from Ager's Omniglor: a guide to writing systems (Ager 1998), in conjunction with Daniels & Bright (1996). Writing systems were chosen from Omniglot graphic writing systems like Chinese and other East Asian writing systems where the character and word levels are not cleanly separable (see Chen et al. 1996; Yeh & Li 2004). When mulvalues indicate BC. The 'phylogeny' column shows the principal sequence of ancestors of the writing system; or when invented, it gives the inventor's name. The fifth column gives the averfigure 4a for more information on stroke-type networks).)

name and sample characters	kind of system	date	phylogeny	avuage character length L	average character munuer of length stroke types L B	characters C	of edges
Ahom W C &	abugida	1250	${ m Brahmi} ightarrow$	2.40	26	40	68
Albanian (Elbasan) d 🕺 🤇	alphabet	1750	invented: unknown	2.60	32	53	82
Ancient Berber (Vertical)	abjad	-150	$\operatorname{Punic} \rightarrow$	2.68	12	25	23
Arabic L + <	abjad	512	Aramaic $ ightarrow$ Nabataean $ ightarrow$	2.63	19	35	36
Arabic Y Y	numeral	512	Aramaic $ ightarrow$ Nabataean Aramaic $ ightarrow$	2.00	10	10	13
Aramaic γ $\frac{1}{2}$	abjad	-900	$Phoenician \rightarrow$	2.68	11	22	25
Armenian (Eastern) 🛄 📙 🦞	alphabet	405	invented: Mesrop-Mashtots	2.21	24	39	42
Asomtavruli U q 1	alphabet	430	Greek	2.42	25	38	52
Avestan w w v	alphabet	250	Aramaic $ ightarrow$ Psalter and Old Pahlavi $ ightarrow$	2.42	35	53	77
Bassa } 🔿 差	alphabet	unknown	invented: unknown	2.37	24	30	38
Batak (Kara Batak) S 7 O	abugida	1350	$Brahmi \rightarrow Pallava \rightarrow Old Kawi \rightarrow$	2.15	13	33	33
Bengali 즉 성 5	abugida	1050	${ m Brahmi} ightarrow { m Devanagari} ightarrow$	3.91	29	45	113
Bengali 🔪 🛫 🧉	numeral	1050	${f Brahmi} o {f D}$ evanagari $ o$	1.80	14	10	15
Brahmir \int_{0} T	abugida	-450	Aramaic \rightarrow	2.14	24	43	51
Buhid (Mangyan) $- \frac{1}{7}$	abugida	1350	$\mathbf{Brahmi} \rightarrow \mathbf{Pallava} \rightarrow \mathbf{Old}\ \mathbf{Kawi} \rightarrow$	3.78	5	18	12
Burmese $\mathcal{O} = \mathcal{O}$	abugida	1150	$\mathbf{Brahmi} \to \mathbf{Mon} \to$	2.51	29	41	85
Burmese J ? ?	numeral	1150	$\mathbf{Brahmi} \to \mathbf{Mon} \to$	1.40	10	10	9
Carrier (Dene)	syllabary (rotating)	1885	invented: Father Adrien-Gabriel Morice	3.20	61	180	68
Celtiberian 🗸 🛃 🕅	alphabet (syllabic)	-550	$\text{Punic} \rightarrow \text{Iberian} \rightarrow$	3.29	6	28	36
Cherokee R 📔 P	syllabary	1819	invented: George Guess	2.35	51	85	102
Chinese	numeral	-1150	unknown	3.00	9	10	13
Cypriot $\uparrow \neq \checkmark$	syllabary	-800	linear $\mathbf{B} o$	3.84	19	55	68
Cyrillic (AbŘhaz) b 3 HO	alphabet	950	$\mathbf{Greek} \to \mathbf{Old} \ \mathbf{Church} \ \mathbf{Slavonic} \to$	3.69	19	62	42
Dehong' a n	abugida	1050	$\mathbf{Brahmi} \to \mathbf{Sinhala} \to$	3.39	13	33	40
Dehong J 2	numeral	1050	$\mathbf{Brahmi} \to \mathbf{Sinhala} \to$	1.90	11	10	12
Deseret $\partial + \beta$	alphabet	1850	invented: George D. Watt	1.68	27	38	34
Devanagari क र्य ग	abugida	1050	$\operatorname{Brahmi} ightarrow$	3.27	30	45	83
Devanagari ? 3	numeral	1050	$\mathbf{Brahmi} \to$	1.40	13	10	×
Dives Akuru 29 2, 29	abugida	1200	$\mathbf{Brahmi} \to \mathbf{Sinhala} \to$	3.77	21	23	31
Enochian X 7 7	alphabet	1580	invented: Dr John Dee and Sir Edward	2.52	19	21	37
Rthinnic (Ga'az) II A L	والتعديطو	350	Kelly Southern I inear Soboen /Minean	7 63	00	Ψ	7.1

Downloaded from http://rspb.royalsocietypublishing.org/ on March 30, 2015

name, and sample characters	kind of system	date	phylogeny	length L	stroke types B	characters <i>C</i>	of edges
Etruscan (archaic) R (E	alphabet	-750	$\operatorname{Greek} \to$	2.91	11	23	25
Faliscan ((((((((((alphabet	-650	$\mathbf{Greek} ightarrow \mathbf{Etruscan} ightarrow$	2.52	15	21	32
Fraser d B f	alphabet	1915	invented: James Ostram Fraser	2.37	17	41	36
Glagolitic 🕂 💾 😗	alphabet	860	$Greek \rightarrow$	4.51	22	41	95
Gothic (Wulfila) A B F	alphabet	350	$Greek \rightarrow$	2.32	17	25	37
Greek $\alpha \beta$ γ	alphabet	-750	$Phoenician \rightarrow$	1.71	21	24	24
Gujarati 5 24 2	abugida	1592	$\mathbf{Brahmi} \rightarrow \mathbf{Devanagari} \rightarrow$	2.21	32	42	63
Gujarati 1 2 3	numeral	1592	${f Brahmi} o {f D}$ evanagari $ o$	1.50	13	10	10
Gurmukhi ब न ट	abugida	1550	$\operatorname{Brahmi} ightarrow$	3.20	31	46	93
Gurmukhi 9 2 3	numeral	1550	$\operatorname{Brahmi} ightarrow$	1.90	12	10	11
Hanuno'o (Mangyan) 7 P V	abugida	1350	$\mathbf{Brahmi} \to \mathbf{Pallava} \to \mathbf{Old}\ \mathbf{Kawi} \to$	3.13	13	16	34
Hebrew 🖸 🎵 🕇	abjad	-125	Aramaic \rightarrow	2.55	18	33	30
Hindu-Arabic 1 2 3	numeral	700	$\mathbf{Brahmi} ightarrow$	1.60	6	10	6
Hungarian Runes 👌 🏌 🅇	alphabet	1000	$\mathbf{Aramaic} \to \mathbf{Sogdian} \to \mathbf{Turkic} \to$	3.08	12	40	34
Hungarian Runes	numeral	1000	$\mathbf{Aramaic} \to \mathbf{Sogdian} \to \mathbf{Turkic} \to$	2.50	4	9	12
Iberian (northern) 🖉 🖡 N	alphabet (syllabic)	-1000	$Punic \rightarrow$	3.46	11	26	27
Iberian (southern) A o Y	alphabet (syllabic)	-1000	$Punic \rightarrow$	3.14	11	22	30
International phonetic J S S	alphabet	1847	invented: Isaac Pitman and Henry Ellis	2.42	47	170	119
Kannada モ シ ビ	abugida	550	$\mathbf{Brahmi} ightarrow$	2.79	34	47	92
Kannada 🔿 💈 🌡	numeral	550	$\mathbf{Brahmi} ightarrow$	1.00	10	10	0
Kharoshthi 🖒 🤇 🖗	abugida	-450	Aramaic \rightarrow	1.72	30	39	44
Kharoshthi) 🎙 🔑	numeral	-450	Aramaic $ ightarrow$	1.88	7	80	11
Khmer R 2 2	abugida	611	${f Brahmi} o {f Pallava} o$	7.49	33	68	70
Khmer g b n	numeral	611	$\mathbf{Brahmi} \to \mathbf{Pallava} \to$	3.70	14	10	30
Korean (Hangeul) - 🖵 🛪	alphabet	1446	invented: King Seycong	2.83	œ	24	19
Kpelle チ ら て	syllabary	1930	invented: Chief Gbili	3.07	74	88	198
Latin, ancient A B C	alphabet	-650	$\mathbf{Greek} \to \mathbf{Etruscan} \to$	2.67	10	21	25
Latin, modern a b c	alphabet	1600	$Greek \rightarrow Etruscan \rightarrow ancient Latin \rightarrow$	2.08	14	26	33
Latin, modern all-caps A B C	alphabet	1600	$\operatorname{Greek} \to \operatorname{Etruscan} \to \operatorname{ancient} \operatorname{Latin} \to$	2.50	17	26	41
Lepcha (Rong) 🐔 🕻 😡	abugida	1720	$\mathbf{Brahmi} \to \mathbf{Devanagari} \to \mathbf{Tibetan} \to$	2.68	44	77	95
Lepcha (Rong) 9 🍣 🍣	numeral	1720	$\mathbf{Brahmi} \to \mathbf{Devanagari} \to \mathbf{Tibetan} \to$	2.60	15	10	27
LimbuZ G &	abugida	1730	$\mathbf{Brahmi} \to \mathbf{Devanagari} \to \mathbf{Tibetan} \to$	2.51	34	37	72
			$\mathbf{Lepcha} \rightarrow$				
Linear B H F	syllabary	-1550	linear $\mathbf{A} ightarrow$	5.03	34	73	148
Marsiliana 🖯 🚡 🗋	alphabet	-650	$Greek \rightarrow$	2.88	15	26	33
Meroitic (non-hieroglyphic)	abugida	-250	Ancient Egyptian $ ightarrow$	3.46	19	23	55
ς ↓ / Messanic∧ R Γ	alnhahet	-550	(-treek	2.87	14	23	34
h Picene) & R		-650	Greek → Etniscan →	2.70	13	23	32
		200	Aramaic →	2.00	16	22	25
، ص		1200	Greek → Asomtavruli →	2.21	30 6	00 7	73
0	aipiiauuu	1400	CICCN - USUILIDAT UI	11.1	2	27	2

Character complexity and redundancy in writing systems M. A. Changizi and S. Shimojo 269

(Continued.)

Table 1. (Continued.)

a glubble glubble 1150 Aramaic - Sogitan - 329 and alphabet 1130 Aramaic - Sogitan - 50 alphabet 1130 Aramaic - Sogitan - 1170 Aramaic - Sogitan - 1170 Aramaic - Sogitan - 1170 Aramaic - 500 alphabet 1030 inverted: Mata Aramai - 1170 Aramaic - 500 alphabet 1040 inverted: Soutemers Katter 2.14 2.0 alphabet 1030 inverted: Soutemers Katter 2.14 2.0 alphabet 1030 inverted: Soutemers Katter 2.14 2.0 alphabet 1030 inverted: Soutemers Katter 2.14 2.0 alphabet 2.	Proc. 1	name, and sample characters	kind of system	date	phylogeny	average character length L	rr number of stroke types B	number of characters <i>C</i>	total number of edges
$\begin{array}{cccccc} & \matcal limit \label{eq:limits} & \matcal limit \label{eq:limit} & \matcal limit eq:l$		Mongolian 🗸 🕹 🕈	alphabet	1150	$\mathrm{Aramaic} \to \mathrm{Sogdian} \to$	3.29	28	35	77
33Nuklakary, $f_{ij} \in f_{ij}$ shid30Armaic1.771873N'Ka, $f_{ij} = 1$ unered1990invented: Maha Arunai1.771874N'Ka, $f_{ij} = 1$ unered1990invented: solutingue Kante2.332.3375Nukha-Muniye79invented: solutingue Kante2.04676Nukha-Muniye79invented: solutingue Kante2.04678Nukha-Muniye79invented: solutingue Kante2.04678Nukha-Muniye79invented: solutingue Kante2.04678Nukha-Muniye79invented: solutingue Kante2.04678Nukha-Muniye39invented: solutingue Kante2.04678Nukha-Muniye39invented: solutingue Kante2.04678Nukha-Muniye39invented: solutingue Kante2.04678Nukha-Muniye39invented: solutingue Kante2.04678Nukha-Muniye310invented: solutingue Kante2.04678Nukha-Muniye3Munike30invented: solutingue Kante2.0479Nukha-Munike3Munike10invented: solutingue Kante2.0479Nukha-Munike3Munike10invented: solutingue Kante2.0479Nukha-Munike3Munike10invented:		Mongolian C D	numeral	1150	Aramaic \rightarrow Sogdian \rightarrow	1.60	11	10	6
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$		Nabataean J S b	abjad	50	Aramaic \rightarrow	1.77	18	22	24
$\begin{aligned} & New Tarlice OI $\ $\ $\ $\ $\ $\ $\ $\ $\ $\ $\ $\ $\ $		Ndjuka' Z F	syllabary	1910	invented: Afaka Atumisi	2.35	36	52	78
$\begin{split} \text{NYCo}(\mathbf{A} + \mathbf{A}) & \text{alphote} & 1949 & \text{invented: Soulemayne Kane 2.14 & 20 \\ \text{NYCo}(\mathbf{A} + \mathbf{A}) & \text{alphote} & 1949 & \text{invented: Soulemayne Kane 2.14 & 20 \\ \text{NYCo}(\mathbf{A} + \mathbf{A}) & \text{alphote} & 18, 0 & \text{alphote} & -650 & \text{Greek} - \text{Brussen} - & 237 & 24 \\ \text{NYCo}(\mathbf{A} + \mathbf{A}) & \text{alphote} & 1390 & \text{invented: Soulemayne Kane 2.14 & 20 \\ \text{Oryard, } & 9(\mathbf{A}) & \text{alphote} & 1390 & \text{invented: Soulemayne Kane 2.14 & 20 \\ \text{Oryard, } & 9(\mathbf{A}) & \text{alphote} & 1390 & \text{invented: Soulemayne Kane 2.13 & 21 \\ \text{Oryard, } & 9(\mathbf{A}) & \text{alphote} & 1390 & \text{invented: Soulemayne Kane 2.14 & 20 \\ \text{Oryard, } & 9(\mathbf{A}) & \text{alphote} & 1390 & \text{invented: Soulemayne Kane 2.13 & 0 \\ \text{Oryard, } & 9(\mathbf{A}) & \text{alphote} & 1390 & \text{invented: Soulemayne Kane 2.24 & 0 \\ \text{Oryard, } & 9(\mathbf{A}) & \text{alphote} & 1091 & \text{invented: Soulemayne Kane 2.24 & 0 \\ \text{Oryard, } & 9(\mathbf{A}) & \text{alphote} & 1091 & \text{invented: Soulemayne Kane 2.24 & 0 \\ \text{Oryard, } & 9(\mathbf{A}) & \text{alphote} & 1091 & \text{invented: Soulemayne Kane 2.24 & 0 \\ \text{Oryard, } & 1000 & \text{invented: Soulemayne Kane 2.24 & 0 \\ \text{Plahon Hilmong U R A & 1000 & \text{invented: Soulemayne Kane 2.24 & 0 & 0 \\ \text{Plahon Hilmong U R A & 1000 & \text{invented: Soulemayne Manne 2.24 & 0 & 0 & 0 \\ \text{Investional A & 1} & alphote (unstaun) 1999 & invented: Soule Law Yang 2.20 & 0 & 0 & 0 & 0 \\ \text{Plahon Hilmong U R A & 1000 & \text{invented: Soulemayne A & 1000 & 0 & 0 & 0 & 0 & 0 \\ \text{Plahon Hilmong U R A & 1000 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 $		New Tai Lue 00 6) 00	numeral	1950	invented: unknown	2.23	12	10	12
NYNCAL F 1 muteral 1949 inverted Soutempter Katte 200 NYNCAL F 1 muteral 1940 inverted Soutempter Katte 200 Organ (Church Strengt 18, Careford 1940 inverted Soutempter Katte 200 Organ (Church Strengt 18, Careford 1940 inverted Soutempter Katte 200 Organ (Church Strengt 18, Careford 1940 inverted Soutempter 1940 inverted Stored Line Yang 2940 inverted Pade Water Antonio 201 inverted Stored Line Yang 2940 inverted Pade Roghmath Muturu 201 inverted Stored Line Yang 2940 inverted Pade Roghmath Muturu 201 inverted Stored Line Yang 2940 inverted Pade Roghmath Muturu 201 inverted Stored Roman 201 inverted Stored Roma	76	N'Ko & 9 H	alphabet	1949	invented: Soulemayne Kante	2.14	20	22	28
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	77	N'Kol P J	numeral	1949	invented: Soulemayne Kante	2.60	9	10	6
Number difference in the second second second second second set is a second set in the second secon	78	North Picene B 8 (alphabet	-650	$\mathbf{Greek} \to \mathbf{Etruscan} \to$	2.67	13	18	27
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	62	Nuskha-khucuri't' IJ J	alphabet	850	$\operatorname{Greek} o \operatorname{Asomtavruli} o$	3.58	16	38	31
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	80	Old Church Slavonic A E B	alphabet	850	$\operatorname{Greek} \to$	3.33	27	42	70
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	81	Old Permic (Abur) Z 🧄 🕁	alphabet	1390	invented: St. Stephen of Perm	3.11	29	38	71
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	82	Oriya କା ଖିଁ ରିଁଁ	abugida	1051	$\mathbf{Brahmi} \to \mathbf{K}\mathbf{alinga} \to$	2.89	34	44	101
$\begin{array}{c c c c c c c c c c c c c c c c c c c $	83	Oriya 🤅 🗿 🎢	numeral	1051	$\mathbf{Brahmi} \to \mathbf{Kalinga} \to$	1.50	15	10	10
Palabarh Himorgi \vec{n} <td>84</td> <td>Oscan A B</td> <td>alphabet</td> <td>-650</td> <td>$Greek \to Etruscan \to$</td> <td>2.71</td> <td>14</td> <td>21</td> <td>28</td>	84	Oscan A B	alphabet	-650	$Greek \to Etruscan \to$	2.71	14	21	28
Palawh Hrnorg ()Talawh Hrnorg ()SImmeral1959invented. Shong Lue Yang2.5015Parthian () Y Palyid (1) Parthian () (2)	85	Pahawh Hmong G M A	alphabet (unusual)	1959	invented: Shong Lue Yang	3.30	42	86	65
Parthim \mathcal{K} $$	86	Pahawh Hmong 🗸 🕃 🚺	numeral	1959	invented: Shong Lue Yang	2.50	15	10	30
Phage-path \mathbf{F}_1 abjad 600 Aramic - Nabataen - Arabic - 2.73 20 Phage-path \mathbf{F}_1 \mathbf{F}_1 abjad 1269 Brahni - Devangaren - Arabic - 2.73 20 Photorian \mathbf{F}_1 \mathbf{F}_1 \mathbf{F}_2 abjad 1269 Brahni - Devangaren - Tibetan - 4.53 211 Photorian \mathbf{F}_1 \mathbf{F}_2 abjad 1209 Brahni - Devangaren - Tibetan - 2.63 100 Redinar (Gananic) \mathbf{F}_2 abjad 130 Brahni - Palara - Old Kawi - 2.00 100 Runic (Danis) Fundar(\mathbf{F}_1 \mathbf{F}_2 abjad -500 bunknown 2.00 100 Runic (Danis) Fundar(\mathbf{F}_1 \mathbf{F}_1 abjad -500 bunknown 2.00 100 Runic (Danis) Fundar(\mathbf{F}_1 \mathbf{F}_1 abjad -500 bunknown 2.03 100 Runic (Danis) Fundar(\mathbf{F}_1 \mathbf{F}_1 abjad -500 bunknown 2.03 100 Sabaean/Minean + \mathbf{F}_1 abjad -500 bunknown 2.03 100 Sanati(O) Cancer) (\mathbf{C}_2 \mathbf{C}_2 \mathbf{R}_1 \mathbf{R}_2 2.00 100 Sanati(O) Cancer) (\mathbf{C}_2 \mathbf{C}_2 \mathbf{R}_2 \mathbf{R}_2 2.03 2.03 Sanati(O) Cancer) (\mathbf{C}_2 \mathbf{C}_2 \mathbf{R}_2 \mathbf{R}_2 2.03 2.03 Sanati(O) Cancer) (\mathbf{C}_2 \mathbf{C}_2 \mathbf{R}_2 2.04 120 Sanati(O) Cancer) (\mathbf{C}_2 \mathbf{C}_2 \mathbf{R}_2 2.04 120 <	87	Parthian 🏷 🗸 🗸	abjad	100	Aramaic \rightarrow	2.59	14	22	25
Thags pathThe standThe stand <td>88</td> <td>ر ، Sashto کر Pashto کر Pashto کر ا</td> <td>abjad</td> <td>600</td> <td>Aramaic \rightarrow Nabataean \rightarrow Arabic \rightarrow</td> <td>2.73</td> <td>20</td> <td>40</td> <td>47</td>	88	ر ، Sashto کر Pashto کر Pashto کر ا	abjad	600	Aramaic \rightarrow Nabataean \rightarrow Arabic \rightarrow	2.73	20	40	47
PhoenicianTheonician	89	لا ما Phags-pa	abugida	1269	$Brahmi \to Devanagari \to Tibetan \to$	4.53	21	40	56
Pollard Miao S3 Δ alphabet (unusual)1905invented: Samuel Pollard2.112.2Paulter J γ γ alphabet (unusual)100 λ ramaic λ ramaic2.0019Runic (Danish Futhark) γ alphabet50unknown2.0310Runic (Danish Futhark) γ alphabet50unknown2.0310Runic (Danish Futhark) γ alphabet50unknown2.0310Runic (Danish Futhark) γ alphabet50unknown2.136Samatian 5 α γ alphabet1920unknown2.135Samatian 5 α γ alphabet1920invented: Pauft Raghunath Murmu3.3329Samati (Ol Camery) δ σ alphabet1920invented: Cismaan Kenadiid1.9029Sil'oti Nagri 7 γ alphabet1920invented: Gama Kenadiid1.9029Sil'oti Nagri 7 γ alphabet1922invented: Gama Kenadiid1.9029Soring Sompeng δ γ alphabet1922invented: Gama Rodi1.9029Soring Sompeng δ γ α α α α 2.2324Soring Sompeng δ γ α α α 2.2924Soring Sompeng δ γ α α α α 2.2929Soring Sompeng δ γ α	06	Phoenician Z	abjad	-1250	northern linear (Canaanite) \rightarrow	2.86	13	22	34
Radret J2abjad100Aramaic -2:0019Runic (Elder Futhark) h h abjad100Brahmi - Old Kawi -2:0010Runic (Elder Futhark) h h alphabet50unknown3:1300Runic (Elder Futhark) h h alphabet50unknown3:1300Runic (Elder Futhark) h h alphabet50unknown3:1300Runic (Elder Futhark) h h alphabet50unknown3:1300Samarian 5 $(G, ener)$ Q g alphabet1920invented: Pandit Raghunath Murmu3:3329Sanali (Ol Cener) Q g alphabet1920invented: Sam Shahalal3:3329Silto in Nagri 7 g q alphabet1920invented: Sam Shahalal3:3329Sanali (Osmanya) S h alphabet1920invented: Sam Shahalal3:3329Somali (Osmanya) S h alphab	91	Pollard Miao S 3 A	alphabet (unusual)	1905	invented: Samuel Pollard	2.11	22	47	32
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	92	Psalter J 3	abjad	100	Aramaic \rightarrow	2.00	19	21	28
Runic (Darish Futhark) \bigwedge \square <td>93</td> <td>Redjang (Kaganga) 🎊 💋 🎢</td> <td>abugida</td> <td>1350</td> <td>$Brahmi \rightarrow Pallava \rightarrow Old Kawi \rightarrow$</td> <td>3.00</td> <td>10</td> <td>36</td> <td>16</td>	93	Redjang (Kaganga) 🎊 💋 🎢	abugida	1350	$Brahmi \rightarrow Pallava \rightarrow Old Kawi \rightarrow$	3.00	10	36	16
Runic (Elder Futharts) Π alphabet50unknown3.136Sabaean/Minean Π <td< td=""><td>94</td><td>Runic (Danish Futhark) 🖗 🆒 Þ</td><td>alphabet</td><td>800</td><td>unknown</td><td>2.63</td><td>10</td><td>16</td><td>21</td></td<>	94	Runic (Danish Futhark) 🖗 🆒 Þ	alphabet	800	unknown	2.63	10	16	21
Sabacan/Minean $H \ge 1$ abjad -500 southern linear \rightarrow Sabacan/Minean \rightarrow 3.55 12Samartian \Im (if \bigtriangledown abjad -500 southern linear \rightarrow Sabacan/Minean \rightarrow 3.55 12Samartian \Im (if \bigtriangledown abjad -500 Aramaic \rightarrow Old Hebrewi $?$ 3.32 29Santali (OI Cemer) \Im (\heartsuit abugida1300invented: Pandir Raghunath Murmu1.4011Somali (OS manya) \circlearrowright \heartsuit abugida1300invented: Saint Raghunath Murmu1.4011Somali (OS manya) \circlearrowright \circlearrowright \backsim abugida1920invented: Saint Raghunath Murmu1.4011Somali (OS manya) \circlearrowright \circlearrowright \backsim abugida1922invented: Saint Shahjalal3.3329Somali (OS manya) \circlearrowright \circlearrowright \backsim abugida1922invented: Ginman Kenadiid1.5012Somali (OS manya) \circlearrowright \circlearrowright \backsim alphaber1922invented: Ginman Kenadiid1.5012Somali (OS manya) \circlearrowright \circlearrowright \circlearrowright alphaber1926invented: Ginman Kenadiid1.9029Sorang Sompeng \circlearrowright \circlearrowright alphaber1926invented: Ginman Kenadiid1.9029Sorang Sompeng \circlearrowright \circlearrowright alphaber1936invented: Mangei Gomango2.2427Sorang Sompeng \circlearrowright \circlearrowright alphaber1936invented: Mangei Gomango1.9012Sorang Sompeng \circlearrowright \circlearrowright alphaber1936invented: Mangei Gomango2.924Sorang Sompeng \circlearrowright \circlearrowright alphaber1936invented: Mangei Gomango2.927Sorang Sompeng \circlearrowright \circlearrowright alphaber1936272817Sorang	95	Runic (Elder Futhark) 🌾 🦒 🕨	alphabet	50	unknown	3.13	9	24	16
Samaritan \mathcal{F} (if ∇ abjad -50 Aramaic -0 ld Hebrew ??? $-$ 4.32 24 Santai (OI Ceneer) \mathcal{D} (G alphabet 1920 invented: Pandit Raghunath Murmu 3.33 22 Santai (OI Ceneer) \mathcal{D} (G alphabet 1920 invented: Pandit Raghunath Murmu 3.33 28 Sunali (OSmanya) \mathcal{E} \mathcal{P} alphabet 1920 invented: Sain Shahjalal 3.33 28 Somali (OSmanya) \mathcal{E} \mathcal{P} alphabet 1920 invented: Sain Shahjalal 3.33 28 Somali (OSmanya) \mathcal{E} \mathcal{P} alphabet 1920 invented: Sain Shahjalal 3.33 28 Somali (OSmanya) \mathcal{E} \mathcal{P} alphabet 1922 invented: Cismaan Kenadiid 1.50 15 Somali (OSmanya) \mathcal{E} \mathcal{P} alphabet 1922 invented: Cismaan Kenadiid 1.50 15 South Arabian \mathcal{E} \mathcal{P} alphabet 1936 invented: Mangei Gomango 1.40 20 South Arabian \mathcal{E} \mathcal{P} abjad -600 south Aramei 2.29 34 Syriacv \mathcal{P} abjad 1686 invented: Mangei Gomango 1.40 12 Syriacv \mathcal{P} abugida 1686 invented: Bogdo Zanabazar 3.63 2.7 Syriacv \mathcal{P} abugida 100 Brahmi \rightarrow Pallava \rightarrow Old Kawi \rightarrow 1.93 17 Tamia \mathcal{P} \mathcal{O} \mathcal{O} invented: Unknown 2.03 29 13 Tamia \mathcal{P} \mathcal{O} \mathcal{O} invented: Unknown 2.03 29 13 Varam Skhih \mathcal{A} \mathcal{A} alphabet 100 Brahmi \rightarrow Pallava \rightarrow Old Kawi \rightarrow 2.74 29 Theban \mathcal{A} \mathcal{A} abugida 1550 invented: Unknown 2.09 23 17 Tamia \mathcal{P} \mathcal{O} \mathcal{O} invented: Unknown 2.09 23 17 South Arama \mathcal{E} \mathcal{O} \mathcal{O} invented: Unknown 2.09 23 17 Tamia \mathcal{P} \mathcal{O} \mathcal{O} invented: Lavo Bodra 2.28 13 Constraint Schih \mathcal{A} \mathcal{A} abugida 150 invented: Lonown 2.09 23 17 South Arama \mathcal{E} \mathcal{O} \mathcal{O} invented: Lonown 2.09 23 17 South Arama \mathcal{E} \mathcal{O} \mathcal{O} invented: Lonown 2.28 13 South Arama \mathcal{E} \mathcal{O} \mathcal{O} invented: Lonown 2.20 23 19 South Arama \mathcal{E} \mathcal{O} \mathcal{O} invented: Lonown 2.20 23 19 South Arama \mathcal{E} \mathcal{O} \mathcal{O} invented: Lonown 2.20 23 19 South Arama \mathcal{E} \mathcal{O} \mathcal{O} invented \mathcal{O} \mathcal	96	Sabaean/Minean 📙 🕺 🤺	abjad	-500	southern linear \rightarrow Sabaean/Minean \rightarrow	3.55	12	29	28
Santali (Ol Cemet') \eth (Ol Cemet') \eth (Ol Cemet') \eth (Ol Cemet') \circlearrowright (S alphabet 1920 invented: Pandit Raghunath Murruu 3.33 29 Santali (Ol Cemet') \circlearrowright \biguplus abugida 1920 invented: Saint Shahjalal 3.33 28 Santali (Osmanya) \circlearrowright \circlearrowright abugida 1920 invented: Cismaan Kenadiid 1.90 29 Sorang Sompeng \circlearrowright \circlearrowright alphabet 1922 invented: Cismaan Kenadiid 1.90 29 Sorang Sompeng \circlearrowright \circlearrowright alphabet 1926 invented: Cismaan Kenadiid 1.90 29 Sorang Sompeng \circlearrowright \circlearrowright alphabet 1936 invented: Cismaan Kenadiid 1.50 15 Sorang Sompeng \circlearrowright \circlearrowright alphabet 1936 invented: Cismaan Kenadiid 1.50 15 Sorang Sompeng \circlearrowright \circlearrowright alphabet 1936 invented: Cismaan Kenadiid 1.50 15 Sorang Sompeng \circlearrowright \circlearrowright alphabet 1936 invented: Mangei Gomango 1.40 12 Sorang Sompeng \circlearrowright \circlearrowright abugida 1036 invented: Mangei Gomango 1.40 229 34 Soronbo \circlearrowright \circlearrowright South Arabian \circlearrowright \circlearrowright abugida 1036 invented: Mangei Gomango 1.40 12 Syncev \checkmark \circlearrowright abugida 100 Brahmi \rightarrow Pallava \rightarrow Old Kawi \leftrightarrow 1.93 17 Tagalog \circlearrowright \circlearrowright abugida 900 Brahmi \rightarrow Pallava \rightarrow Old Kawi \leftrightarrow 2.27 2.27 2.27 Tagalog \circlearrowright \circlearrowright abugida 150 Brahmi \rightarrow Pallava \rightarrow Old Kawi \leftrightarrow 2.29 13 Theban \circlearrowright \circlearrowright abugida 150 Brahmi \rightarrow Pallava \rightarrow Old Kawi \leftrightarrow 2.28 13 Theban \circlearrowright \circlearrowright abugida 150 invented: Unknown 3.83 2.86 17 Varam z Kaing \circlearrowright \circlearrowright alphabet 100 Punic \rightarrow Ancient Berber \rightarrow 2.86 17	26	Samaritan 🕉 🕼 🔻	abjad	-50	Aramaic $ ightarrow$ Old Hebrew ??? $ ightarrow$	4.32	24	22	72
Santali (Ol Cemet') $\langle \mathbf{v} \in \mathbf{e}$ numeral 1920 invented: Pandit Raghunath Murmu 1.40 111 Somali (Osmanya) $\langle \mathbf{v} : \mathbf{v} = \mathbf{e}$ abugida 1300 invented: Saint Shahjalal 3.33 28 Sorang Sompeng $\langle \mathbf{v} : \mathbf{v} = \mathbf{e}$ alphabet 1922 invented: Cismaan Kenadiid 1.50 15 Sorang Sompeng $\langle \mathbf{v} : \mathbf{v} = \mathbf{e}$ alphabet 1922 invented: Cismaan Kenadiid 1.50 15 Sorang Sompeng $\langle \mathbf{v} : \mathbf{v} = \mathbf{e}$ alphabet 1936 invented: Cismaan Kenadiid 1.50 15 Sorang Sompeng $\langle \mathbf{v} : \mathbf{v} = \mathbf{e}$ alphabet 1936 invented: Saint Shahjalal 1.50 15 Sorang Sompeng $\langle \mathbf{v} : \mathbf{v} = \mathbf{e}$ alphabet 1936 invented: Saint Shahjalal 1.40 12 Sorang Sompeng $\langle \mathbf{v} : \mathbf{v} = \mathbf{e}$ alphabet 1936 invented: Saint Somango 1.40 12 Sorang Sompeng $\langle \mathbf{v} : \mathbf{v} = \mathbf{e}$ alphabet 1936 invented: Mangei Gomango 1.40 12 Sorond Sompeng $\langle \mathbf{v} : \mathbf{v} = \mathbf{e}$ alphabet 1936 invented: Bogo Zanabazar 3.63 2.27 2.57 Tagalog $\langle \mathbf{v} : \mathbf{v} = \mathbf{e}$ albidid 200 Brahmi $\rightarrow \operatorname{Pallava} \rightarrow \operatorname{Old} Kawi \rightarrow 2.27$ 2.27 2.57 Tagabawa $\langle \mathbf{v} : \mathbf{v} = \mathbf{e}$ abugida 1550 invented: Unknown 2.03 Theban $\langle \mathbf{q} : \mathbf{m} = \operatorname{abugida} -300$ Brahmi $\rightarrow \operatorname{Pallava} \rightarrow \operatorname{Old} Kawi \rightarrow 2.74$ 29 Thaana $\langle \mathbf{r} : \mathbf{p} = \operatorname{abugida} -300$ Brahmi $\rightarrow \operatorname{Pallava} \rightarrow \operatorname{Old} Kawi \rightarrow 2.23$ 19 Thaana $\langle \mathbf{r} : \mathbf{r} = \operatorname{abugida} -300$ Brahmi $\rightarrow \operatorname{Pallava} \rightarrow \operatorname{Old} Kawi \rightarrow 2.23$ 19 Thaana $\langle \mathbf{r} : \mathbf{r} = \operatorname{abugida} -300$ Brahmi $\rightarrow \operatorname{Pallava} \rightarrow \operatorname{Old} Kawi \rightarrow 2.23$ 19 Thaana $\langle \mathbf{r} : \mathbf{r} = \operatorname{abugida} -300$ Brahmi $\rightarrow \operatorname{Pallava} \rightarrow \operatorname{Old} Kawi \rightarrow 2.23$ 19 Thaana $\langle \mathbf{r} : \mathbf{r} = \operatorname{abugida} -300$ Brahmi $\rightarrow 2.24$ 20 Thifinagh $\langle \mathbf{r} : \mathbf{r} = \operatorname{abugida} -100$ Punic $\rightarrow \operatorname{Ancient Berber} \rightarrow 2.48$ 16 Thriang $\langle \mathbf{r} : \mathbf{r} = \operatorname{abugida} -100$ Three the Brahmi $\rightarrow 2.86$ 17 Three A muricum Berber $\rightarrow 2.48$ 16 Three A muricum A muricum Berber $\rightarrow 2.48$ 16 Three A	98	Santali (Ol Cemet') 🐉 🜔 G	alphabet	1920	invented: Pandit Raghunath Murmu	3.33	29	30	80
Sil'oti Nagri $\mathcal{F}_{\mathcal{F}}$ $\mathcal{F}_{\mathcal{F}}$ abugida 1300 invented: Saint Shahjalal 3.33 28 Somali (Osmanya) $\mathcal{E}_{\mathcal{F}}$ $\mathcal{F}_{\mathcal{F}}$ alphabet 1922 invented: Cismaan Kenadiid 1.90 29 Somali (Osmanya) $\mathcal{E}_{\mathcal{F}}$ $\mathcal{F}_{\mathcal{F}}$ alphabet 1922 invented: Cismaan Kenadiid 1.50 15 Sorang Sompeng $\mathcal{F}_{\mathcal{F}}$ alphabet 1936 invented: Mangei Gomango 2.29 34 Sorang Sompeng $\mathcal{F}_{\mathcal{F}}$ alphabet 1936 invented: Mangei Gomango 2.29 34 Sorang Sompeng $\mathcal{F}_{\mathcal{F}}$ abjad -600 suthern linear \rightarrow 3.29 13 Soyombo $\mathcal{F}_{\mathcal{F}}$ abjad 400 Brahmi \rightarrow Pallava \rightarrow Old Kawi \rightarrow 2.27 25 Tagalog \mathcal{O} \mathcal{F} \mathcal{F} abugida 900 Brahmi \rightarrow Pallava \rightarrow Old Kawi \rightarrow 2.27 25 Theban $\mathcal{F}_{\mathcal{F}}$ abugida 900 Brahmi \rightarrow Pallava \rightarrow Old Kawi \rightarrow 2.27 25 Theban $\mathcal{F}_{\mathcal{F}}$ abugida 1560 invented: Unknown 2.23 19 Theban $\mathcal{F}_{\mathcal{F}}$ abugida 1500 Brahmi \rightarrow Pallava \rightarrow Old Kawi \rightarrow 2.27 25 Theban $\mathcal{F}_{\mathcal{F}}$ abugida 1500 Brahmi \rightarrow Pallava \rightarrow Old Kawi \rightarrow 2.24 29 Theban $\mathcal{F}_{\mathcal{F}}$ abugida 1500 Brahmi \rightarrow Pallava \rightarrow Old Kawi \rightarrow 2.24 29 Theban $\mathcal{F}_{\mathcal{F}}$ alphabet 100 brahmi \rightarrow Pallava \rightarrow Old Kawi \rightarrow 2.24 29 Theban $\mathcal{F}_{\mathcal{F}}$ alphabet 100 brahmi \rightarrow 2.24 209 23 Theban $\mathcal{F}_{\mathcal{F}}$ alphabet 100 brahmi \rightarrow Pallava \rightarrow Old Kawi \rightarrow 2.24 10 Varanz Kshif $\mathcal{F}_{\mathcal{F}}$ alphabet 100 invented: Unknown 2.09 23 Theban $\mathcal{F}_{\mathcal{F}}$ alphabet 200 brahmi \rightarrow 2.48 16	66	Santali (Ol Cemet') 🖓 🖻 😢	numeral	1920	invented: Pandit Raghunath Murmu	1.40	11	10	7
Somali (Osmanya) $\delta_{1} = \frac{1}{7}$, alphabet 1922 invented: Cismaan Kenadiid 1.90 29 Sorang Sompeng $\delta_{1} = \frac{1}{7}$, numeral 1922 invented: Cismaan Kenadiid 1.50 15 Sorang Sompeng $\delta_{2} = \frac{1}{7}$, numeral 1922 invented: Cismaan Kenadiid 1.50 15 Sorang Sompeng $\delta_{2} = \frac{1}{7}$, numeral 1936 invented: Mangei Gomango 2.29 34 South Arabian $\delta_{2} = \frac{1}{7}$, alphabet 1936 invented: Mangei Gomango 2.29 34 South Arabian $\delta_{2} = \frac{1}{7}$, alphabet 1936 invented: Mangei Gomango 2.29 34 South Arabian $\delta_{2} = \frac{1}{7}$, abugida 1686 invented: Mangei Gomango 2.29 37 Syriacv $\sim \delta_{2} = \frac{1}{7}$ abugida 1686 invented: Mangei Gomango 2.29 37 Syriacv $\sim \delta_{2} = \frac{1}{7}$ abugida 900 Brahmi $\rightarrow Pallava \rightarrow Old Kawi \rightarrow 2.27$ 25 Tamil $\mathcal{G}_{1} = \frac{1}{7}$, abugida 900 Brahmi $\rightarrow Pallava \rightarrow Old Kawi \rightarrow 2.23$ 19 Theban $\mathcal{H}_{2} = \frac{1}{7}$ abugida 150 invented: Unknown 2.03 23 Theban $\mathcal{H}_{2} = \frac{1}{7}$ alphabet 100 invented: Unknown 2.09 23 Varang Kshiri $\beta_{1} = \frac{1}{7}$ alphabet 100 invented: Unknown 2.08 Arabia 28 Varang Kshiri $\beta_{1} = \frac{1}{7}$ alphabet 100 invented: Luknown 2.88 116 Varang Kshiri $\beta_{1} = \frac{1}{7}$ alphabet 1900 invented: Luknown 2.88 116	100	Sil'oti Nagri 7 7 ज	abugida	1300	invented: Saint Shahjalal	3.33	28	34	69
Somali (Osmanya) $\xi \in \pi$ numeral 1922 invented: Cismaan Kenadiid 1.50 15 Sorang Sompeng $\xi \in \pi$ numeral 1936 invented: Mangei Gomango 2.29 34 Sorang Sompeng $\xi \in \pi$ alphabet 1936 invented: Mangei Gomango 2.29 34 Sorang Somport $\Im \Im \Im \Im \Im $ abjad -600 southern linear \rightarrow 3.29 113 Soyombot $\Im \Im \Im \Im $ abjad -600 southern linear \rightarrow 3.29 113 Syriacv, $\Rightarrow \checkmark$ abjad 400 Brahmi \rightarrow 1008 invented: Boglo Zanabazar 3.63 27 Syriacv, $\Rightarrow \circlearrowright$ abugida 900 Brahmi \rightarrow Pallava \rightarrow Old Kawi \rightarrow 1.93 17 Tagalog $\Sigma \mathcal{K} \mathcal{K}$ abugida 900 Brahmi \rightarrow Pallava \rightarrow Old Kawi \rightarrow 2.27 25 Thebant $\Im \Im$ abugida 1550 invented: Unknown 3.83 28 Umbriang $\Im \Im$ alphabet -100 Punic \rightarrow Ancient Berber \rightarrow 2.48 16 Varang Kshiti $\Im \Lambda$ \Im alphabet -350 Greek \rightarrow Euroson \rightarrow 2.48 16	101	Somali (Osmanya) 🗗 🛂 🤸	alphabet	1922	invented: Cismaan Kenadiid	1.90	29	30	38
Sorang Sompeng $\left(\begin{array}{cccc} 3 \\ 3 \\ 3 \\ 3 \\ 3 \\ 3 \\ 3 \\ 3 \\ 3 \\ 3 $	102	Somali (Osmanya) f E Ā	numeral	1922	invented: Cismaan Kenadiid	1.50	15	10	10
Sorang Sompeng $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ numeral 1936 invented: Mangei Gomango 1.40 12 South Arabian $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ abjad -600 southern linear \rightarrow 3.29 13 Syriac $\begin{pmatrix} 1 \\ 3 \end{bmatrix}$ abjad -600 invented: Bogdo Zanabazar 3.63 27 Syriac $\begin{pmatrix} 1 \\ 3 \end{bmatrix}$ abjad 400 Aramaic \rightarrow 2.27 25 Tagalog $\begin{pmatrix} 1 \\ 3 \end{bmatrix}$ abugida 900 Brahmi \rightarrow Pallava \rightarrow Old Kawi \rightarrow 1.93 17 Tamil $\begin{bmatrix} 6 \\ 3 \\ 3 \end{bmatrix}$ abugida -300 Brahmi \rightarrow Pallava \rightarrow Old Kawi \rightarrow 2.23 19 Theban $\begin{bmatrix} 1 \\ 4 \\ 1 \end{bmatrix}$ abugida -300 Brahmi \rightarrow Pallava \rightarrow Old Kawi \rightarrow 2.23 29 Theban $\begin{bmatrix} 1 \\ 4 \\ 1 \end{bmatrix}$ abugida -300 Brahmi \rightarrow Pallava \rightarrow Old Kawi \rightarrow 2.23 19 Threata $\begin{bmatrix} 5 \\ 3 \\ 1 \end{bmatrix}$ abugida -300 Brahmi \rightarrow 2.74 29 Threata $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ abugida -300 Brahmi \rightarrow Pallava \rightarrow Old Kawi \rightarrow 2.24 29 Threata $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ abugida -300 Brahmi \rightarrow 2.74 29 Trifinagh \mp 1 $\begin{pmatrix} 3 \\ 4 \end{bmatrix}$ alphabet 100 Punic \rightarrow Ancient Berber \rightarrow 2.48 16 Invented: Lako Bodra 2.86 17 The bank t 1900 invented: Lako Bodra 2.86 17	103	Sorang Sompeng р 🐧 🖥	alphabet	1936	invented: Mangei Gomango	2.29	34	24	58
South Arabian § β	104	Sorang Sompeng	numeral	1936	invented: Mangei Gomango	1.40	12	10	8
Soyomboff \overline{A} abugida1686invented: Bogdo Zanabazar3.6327Syriac $\nabla_{\mathbf{v}}$ \overline{A} abugida400Aramaic \rightarrow 2.2725Syriac $\nabla_{\mathbf{v}}$ abugida900Brahmi \rightarrow Pallava \rightarrow Old Kawi \rightarrow 1.9317Tagbanwa \times \widetilde{E} abugida900Brahmi \rightarrow Pallava \rightarrow Old Kawi \rightarrow 2.2725Tamil \mathcal{G} \widetilde{O} abugida-300Brahmi \rightarrow Pallava \rightarrow Old Kawi \rightarrow 2.2319Thaana \checkmark \checkmark abugida1550invented: Unknown2.0923Theban \mathcal{H} alphabet100invented: Unknown3.8328Tifinagh \ddagger \Rightarrow alphabet100Punic \rightarrow Ancient Berber \rightarrow 2.4816Varang Kshiti \mathcal{A} alphabet1900invented: Lako Bodra2.4816	105	South Arabian 8 8 4	abjad	-600	southern linear \rightarrow	3.29	13	28	36
Syriac V L abjad400Aramaic \rightarrow 2.2725Tagalog \mathcal{O} \mathcal{K} abugida900Brahmi \rightarrow Pallava \rightarrow Old Kawi \rightarrow 1.9317Tagbanwa \times \mathcal{K} abugida900Brahmi \rightarrow Pallava \rightarrow Old Kawi \rightarrow 2.2319Tamil \mathcal{G} \mathcal{O} abugida -300 Brahmi \rightarrow Pallava \rightarrow Old Kawi \rightarrow 2.2319Thaana \mathcal{I} abugida -300 Brahmi \rightarrow 2.74 29Theban \mathcal{H} \mathcal{I} alphabet100invented: Unknown2.0923Tifinagh \ddagger \mathcal{I} alphabet100Punic \rightarrow Ancient Berber \rightarrow 2.8813Varang Kshiti \mathcal{I} \mathcal{I} alphabet1900invented: Lako Bodra2.4816	106	Soyombo H 3 X	abugida	1686	invented: Bogdo Zanabazar	3.63	27	35	88
Tagalog Tagbanwa Tagbanwa Tamil HTagalog S S STabugida900 Brahmi \rightarrow Pallava \rightarrow Old Kawi \rightarrow 1.9317Tagbanwa Tagbanwa Tamil HG) C) C)abugida abugida900 Brahmi \rightarrow Pallava \rightarrow Old Kawi \rightarrow 2.2319Tamil HG) C) Thaana Theban \mathfrak{H} abugida a bugida-300 Brahmi \rightarrow Brahmi \rightarrow 2.2319Thaana Theban \mathfrak{H} Tifinagh \ddagger 1550 alphabetinvented: Unknown invented: Unknown2.0923Tifinagh \ddagger Tifinagh \ddagger 2alphabet abjad-100 -100Punic \rightarrow Ancient Berber \rightarrow 2.8813Varang Kshiti A 4 alphabet1900 invented: Lako Bodra2.4816	107	Syriac v L	abjad	400	Aramaic \rightarrow	2.27	25	22	50
Tagbanwa X \mathcal{F} abugida900Brahmi \rightarrow Pallava \rightarrow Old Kawi \rightarrow 2.2319Tamil \mathcal{G} (\mathcal{G}) \mathcal{G})abugida -300 Brahmi \rightarrow 2.74 29Thaana \mathcal{F} \mathcal{P} abugida1550invented: Unknown2.0923Theban \mathcal{H} \mathcal{A} abugida -150 invented: Unknown2.0923Tifinagh \ddagger T \mathcal{K} abjad -100 Punic \rightarrow Ancient Berber2.8813Varang Kshiti \mathcal{A} \mathcal{A} \mathcal{A} alphabet -350 Greek \rightarrow Erruscan \rightarrow 2.4816Varang Kshiti \mathcal{A} \mathcal{A} \mathcal{A} \mathcal{A} \mathcal{A} \mathcal{A} \mathcal{A} \mathcal{A}	108	Tagalog L L	abugida	006	$\text{Brahmi} \rightarrow \text{Pallava} \rightarrow \text{Old Kawi} \rightarrow$	1.93	17	16	23
Tamil \mathcal{G} <t< td=""><td>109</td><td>Tagbanwa X 😚 🏎</td><td>abugida</td><td>006</td><td>ightarrow Pallava -</td><td>2.23</td><td>19</td><td>15</td><td>26</td></t<>	109	Tagbanwa X 😚 🏎	abugida	006	ightarrow Pallava -	2.23	19	15	26
Thaana \checkmark \checkmark abugida1550invented: Unknown2.0923Theban \varUpsilon \varUpsilon alphabet100invented: Unknown3.8328Tifinagh \ddagger \updownarrow alphabet100punic \rightarrow Ancient Berber2.8813Umbriang \Im alphabet -350 Greek \rightarrow Etruscan \rightarrow 2.4816Varang Kshiti \varUpsilon Π alphabet1900invented: Lako Bodra2.8617	110		abugida	-300		2.74	29	34	62
Theban \mathcal{A}_{1} alphabet100invented: Unknown3.8328Tifinagh $\mathbf{\pm} 1 \mathbf{\times}$ abjad -100 Punic \rightarrow Ancient Berber \rightarrow 2.8813Umbrian \mathcal{A} \mathcal{A} \mathcal{A} alphabet -350 Greek \rightarrow Etruscan \rightarrow 2.4816Varane Kshit \mathcal{B} \mathcal{A} \mathcal{A} alphabet1900invented: Lako Bodra2.8617	111	Thana C	abugida	1550	invented: Unknown	2.09	23	35	42
Tifinagh # I $\ddot{\mathbf{X}}$ abjad -100 Punic \rightarrow Ancient Berber2.8813Umbriang \mathbf{g} alphabet -350 Greek \rightarrow Etruscan \rightarrow 2.4816Varang Kshiti \mathbf{S} \mathbf{d} \mathbf{q} alphabet1900invented: Lako Bodra2.8617	112	Theban Y d W	alphabet	100	invented: Unknown	3.83	28	24	69
Umbrian β $\widehat{3}$ alphabet -350 Greek \rightarrow Etruscan \rightarrow 2.48 16 Varang Kshiti \mathfrak{H} \mathfrak{A} \mathfrak{H} alphabet 1900 invented: Lako Bodra 2.86 17	113	Tifinagh # ⊢ ¥	abjad	-100	$Punic \rightarrow Ancient Berber \rightarrow$	2.88	13	25	23
Varang Kshiti 2 4 alphabet 1900 invented: Lako Bodra 2.86 17	114	Umbriand A 3	alphabet	-350	Greek ightarrow Etruscan ightarrow	2.48	16	21	30
	115	Varang Kshiti 2 1 1	alphabet	1900	invented: Lako Bodra	2.86	17	21	26

Downloaded from http://rspb.royalsocietypublishing.org/ on March 30, 2015

Table 1. (Continued.)

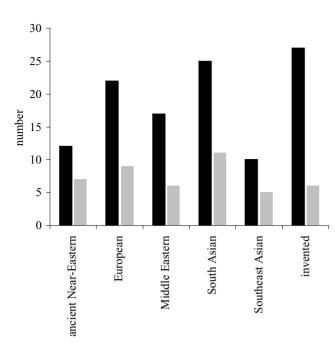


Figure 1. Distribution of writing systems used in the study (see table 1) across the major phylogenetic classes (black bars), and also the number of sections devoted to the phylogenetic classes in Daniels & Bright (1996) *The world's writing systems*, the most exhaustive book on the topic (grey bars). Among the major phylogenetic classes, the distributions are highly correlated ($r^2 = 0.81$).

and jointly determine the third via the equation R = 1 - (d/L). There are two qualitatively different ways that strokes could be combinatorially used to build characters (Changizi 2001, 2003*a*,*b*). The first is the *universal stroke-type approach*, where the number of stroke types does not vary as a function of writing system size, and greater numbers of characters. The second is the *invariant-length approach*, where character length is invariant, and greater numbers of characters are accommodated by increasing the number of stroke types from which characters are built.

To test whether either of these two scaling approaches applies to writing systems, we measured average character lengths for all 115 writing systems in table 1 (see figure 1 for the distribution of classes of writing system). Figure 2aillustrates the manner in which characters are decomposed into strokes. Our first result is that writing systems appear to conform to the invariant-length approach (see figure 2b), with average lengths of approximately 3 (but with lower lengths of approximately 2 for number systems).

For the remainder of § 2, we describe two distinct methods for estimating the combinatorial degree, $d = \beta L$. From the length, *L*, and combinatorial degree, *d*, we will be able to compute the redundancy, R = 1 - (d/L).

The first method of estimating combinatorial degree is to plot stroke-type repertoire size, *B*, as a function of writing system size, *C*, and measure the scaling exponent. Recall that $C \propto B^d$. Thus, $B \propto C^{1/d}$, and the best-fit slope on a loglog plot of *B* versus *C* is an estimate of 1/*d*. Figure 3*a* illustrates the manner in which the stroke-type repertoire is determined, and figure 3*b* describes tests of repeatability. We let d_{BC} denote estimates of the combinatorial degree via this first method. We report four estimates from the data, employing four different subscripts to distinguish between them-'all', 'alpha', 'all,bin' and 'alpha,bin'-where the subscript 'all' means that all writing systems from table 1 are used in the estimate, 'alpha' means that only the nonnumber systems are used, and 'bin' means that the estimate is taken from a binned plot. Figure 3c shows the plot of all the data, and $B \propto C^{0.63}$, and thus $d_{BC,all} = 1/0.63 = 1.60$. The inset of figure 3c shows the binned version of all the data, and $d_{BC,all,bin} = 1.49$. Excluding numerals, the respective estimates are $d_{BC,alpha} = 1.36$ and $d_{BC,alpha,bin} =$ 1.33. These combinatorial degree estimates therefore range from 1.33 to 1.60. Given the average character lengths, the redundancy corresponding estimates four are: $R_{BC,all} = 41\%, R_{BC,all,bin} = 46\%, R_{BC,alpha} = 53\%$ and $R_{BC,alpha,bin} = 56\%$.

The second method for determining the combinatorial degree is via measuring how stroke-type 'interactiveness' changes with writing system size. We let d_{deg} denote estimates of the combinatorial degree via this second method. If strokes are used combinatorially, then in a larger writing system, any given stroke type must, on average, be able to interact with a greater number of stroke types. The *degree*, δ , of a stroke type is the total number of stroke-types with which the stroke type intersects, across all characters of the writing system (for cases of unconnected strokes, like the dot of an 'i', the stroke was deemed connected to the nearest stroke). Figure 4a illustrates the manner in which the stroke-type degree is determined. Figure 4b shows that the average stroke-type degree changes slowly with writing system size, consistent with a power law $\delta \propto C^{w}$: Scaling exponents via the four methods are $w_{all} = 0.24$, $w_{all,bin} = 0.17$, $w_{alpha} = 0.13$, $w_{alpha,bin} = 0.10$. From this it is possible to compute the combinatorial degree, as we now explain. How many characters can be built with length L? In writing a single character, there are B stroke types one may begin with, and for each of these there are, on average, δ many stroke types which may be drawn next, and for each of these, δ more, and so on until all L strokes have occurred in the character. Thus, the number of characters that can be built is $C = B\delta^{L-1}$. Given that $\delta \propto C^{w}$, we can write $C \propto$ $B(C^w)^{L-1}$, and solving for C, we have that $C \propto B^{1/[1-w(L-1)]}$. Therefore, the combinatorial degree can be estimated as $d_{\text{deg}} = 1/[1 - w(L-1)]$, where w is the scaling exponent for stroke-type degree as a function of writing system size (see Changizi et al. (2002) for related observations). The four estimates of combinatorial degree using stroke-type degree scaling are $d_{\text{deg,all}} = 1.73$, $d_{\text{deg,all,bin}} = 1.42$, $d_{\text{deg,alpha}} = 1.32$ and $d_{\text{deg,alpha,bin}} = 1.26$. These combinatorial degree estimates therefore range from 1.26 to 1.73, similar to the range of 1.33 to 1.60 that we found earlier via the first method. The corresponding redundancies are $R_{\text{deg,all}} = 36\%$, $R_{\text{deg,all,bin}} = 49\%$, $R_{\text{deg,alpha}} = 54\%$ and $R_{\text{deg,alpha,bin}} =$ 58%.

3. DISCUSSION

We found that writing systems have average character lengths of approximately 3 (number systems being the exception, with an average of approximately 2). And, via two distinct kinds of measurement and analysis, we found that the combinatorial degrees for writing systems are very approximately 3/2, and redundancies *ca*. 50%. Importantly, these values appear to not much vary as a function of writing system size. Because the combinatorial degree is

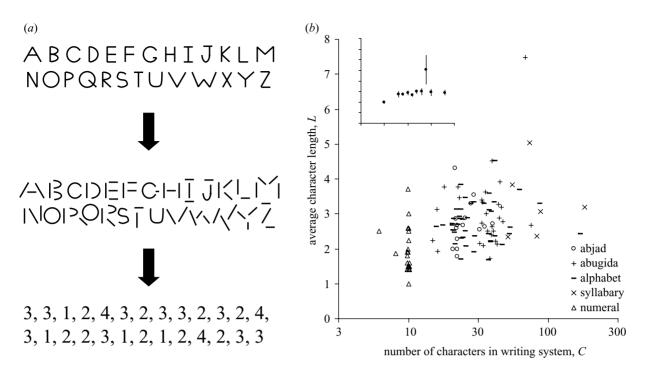


Figure 2. (a) Illustration of the method for determining character lengths (i.e. the number of strokes per character). Each character is decomposed into separable strokes, where strokes are separated by discontinuities so that 'U' is one stroke but 'V' is two, and also stroke junctions are decomposed into their constituents so that 'T' and 'X' junctions possess two strokes, 'Y', 'K' and ' Ψ ' junctions possess three strokes, etc. Three naïve observers were asked to decompose characters into strokes, and there were no disagreements. (b) Plot of average character length versus the number of characters (on a log scale), for 115 writing systems. Data are labelled by abjad (characters for consonants but not vowels), abugidas (characters for consonants and diacritic symbols for vowels), alphabets (characters for consonants and vowels), syllabaries (characters for syllables such as 'ba', 'be', 'bi', etc.) and numerals (characters for numbers). *x*-axis values have been randomly perturbed by $\pm 1\%$ to help distinguish the points on the plot. The average length is 2.79 for invented systems (a set of 38 *independent* writing systems) and 2.70 for non-invented systems. Inset: plot of the same data, and same axes, but average character lengths binned at 0.1 intervals along the *x*-axis (standard error bars shown). One can see that, except for number systems where the average length is approximately 2 (the average across the average lengths of the 22 numeral systems is 1.95, with standard error 0.14), the average character length does not appear to vary as a function of writing system size (the average across the average lengths of the 93 non-numeral systems is 2.91, with standard error 0.09). These data mean that human writing systems conform to the invariant-length approach to accommodating writing systems of greater size.

significantly above 1, it means that writing systems use strokes in a genuinely combinatorial fashion (i.e. doubling the number of strokes more than doubles the number of characters, because $C \propto B^{3/2}$). However, although writing systems are combinatorial, they are not very combinatorial, because the combinatorial degree of 3/2 is not much greater than 1. Because average character lengths are approximately 3, the maximum possible combinatorial degree is 3. Because only approximately half of the total possible degrees of freedom is used, the redundancy is ca. 50%. Characters therefore tend to be about twice as long (in number of strokes) as they need to be. Alternatively, the combinatorial degree could be twice what it is, which would allow the number of stroke types to grow much more slowly as a function of writing system size (namely as the cube root) than they in fact do. These results may have implications for the future design of writing systems.

4. CONCLUSION

Writing systems are under selective pressure to be easy to read and write, but there are reasons to think that the principal pressure is for ease of reading. First, text is written only once, whereas it may be read arbitrarily many times. The utilities due to reading will accordingly be amplified

relative to that for writing. Many writing systems throughout history, however, were not read to the extent that contemporary writing systems are, and this argument will not apply as strongly to such writing systems. Second, cursive scripts and shorthand are two classes of writing system where selection is primarily driven by writing optimization, and in these cases the characters are qualitatively very different compared with those of the typical writing system, and are more difficult to read. Third, and last, typeface and computer fonts are two classes of script where there is no selective pressure for writing at all, and characters in these scripts are qualitatively quite similar to those of the typical writing system. None of these above arguments alone is strong, but together they give us some reason to suspect that the principal selective pressure on most writing systems may come from vision. Assuming this, we ask, is there something about these fundamental properties of writing systems that might be 'good' for the visual system?

Consider redundancy first. Because character recognition requires recognizing the strokes (Pelli *et al.* 2004), and because strokes tend have small angular size and high shape variability, some redundancy is useful so that misrecognition of one or two strokes does not necessarily lead to misrecognition of the character. Why should the visual

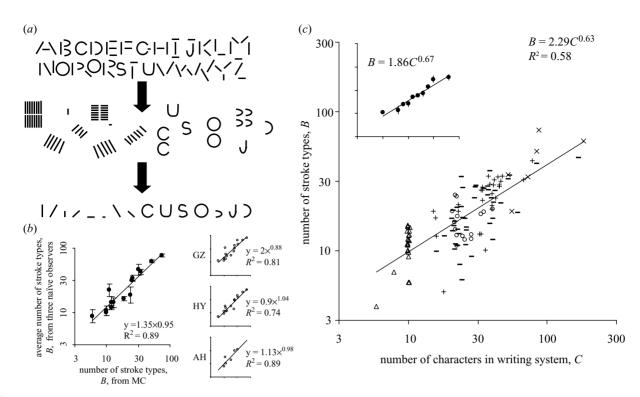


Figure 3. (a) Illustration of how the stroke-type repertoire is determined for a writing system. After the characters are decomposed into their constituent strokes (see figure 1a), the strokes are clustered near strokes that appear to be similar. Stroke types were determined by the primary author (M.C.) on the basis of high intra-cluster similarity in orientation, shape and length. (b) As a test of repeatability, three naïve observers (G.Z., H.Y. and A.H.) were asked to determine the stroke-type repertoire for a wide variety of writing systems (G.Z. and H.Y. carried this out for Ancient Berber, Ahom, Albanian, Arabic, Arabic numerals, Aramaic, Armenian, Asomtavruli, Avestan, Hanuno'o, Cherokee, Hungarian Runes, Elder Futhark, Danish Futhark, Kpelle; and A.H. carried this out for just the first six). On the left is a log-log plot of the average stroke-type repertoire size measured by the three naïve observers versus the estimates of M.C. Standard error bars are shown, as well as the best-fit (by linear regression) equation and line, and the correlation. One can see that the correlation is high and that the exponent relating them is approximately 1, meaning that naïve observers' estimates of stroke-type repertoire size scale in direct proportion to the estimates of M.C. The three plots on the right possess the same x-axis as the one on the left, but the y-axis now has each individual observer's stroke-type estimates. The effects of systematic under- or over-counting (as seen for example in G.Z.) will affect the proportionality constant relating stroke-type repertoire size, B, to writing system size, C, but not the scaling exponent, which is what is of interest to us here. (c) Plot of number of stroke types versus number of characters for 115 writing systems. Circles, abjad; plus symbols, abugida; minus symbols, alphabet; crosses, syllabary; and triangles, numerical. The linear regression line and equation are shown, along with correlation. Data points on each axis have been perturbed by $\pm 1\%$ to aid in their discrimination. The best-fit relationship is $B = 3.18C^{0.57}$ for invented systems (a set of independent data), and $B = 2.31C^{0.60}$ for non-invented systems. Inset: same plot, and same axes, but stroke-type repertoire sizes binned at 0.1 intervals along the log C-axis (standard error bars shown).

system prefer character lengths of approximately 3? The value of 3 naturally suggests the 'subitizing limit', which is the number of objects that can be stored in visual short term memory, and is often put at roughly 3 (e.g. Trick & Pylyshyn 1994; Vogel et al. 2001). That is, perhaps there are, on average, three strokes per character, independent of writing system size, because all the strokes can be simultaneously processed, whereas processing times increase substantially for greater than around three objects. It has been thought that this may underly why number systems tend to represent '1' by one stroke, '2' by two strokes, and '3' by three strokes, but this stops for greater numbers (Ifrah 1985; Zhang & Norman 1995; Dehaene 1997). The combinatorial degree value of 3/2, and the connected rate at which the number of stroke types increases with writing system size (namely as the 3/2 power), would be a consequence of the redundancy and subitizing limit.

A distinct possible kind of explanation for the average length of 3 concerns bottom-up, hierarchical processing of

characters (A. Hampton, private communication). Imagine a lower-level retinotopic map in visual cortex, where the L strokes of a character are simultaneously recognized in L nearby regions of the cortex. Suppose also that multiple nearby regions in the lower level connect in a feed-forward fashion to a single region of the upper-level retinotopic area. A single upper-level region could integrate only from as many lower-level regions as connect to it, and perhaps character length L would be constrained by this. Supposing that multiple lower-level regions feed-forward to an upper-level region if and only if the lower-level regions are all mutually adjacent, and assuming that regions are hexagonally packed in neocortex, each upperlevel region will integrate exactly three mutually adjacent lower-level regions, which would cohere with the average length of $L \approx 3$.

Another intriguing possibility is that there is a fundamental ecological explanation for these writing system features. The visual system has been selected to quickly

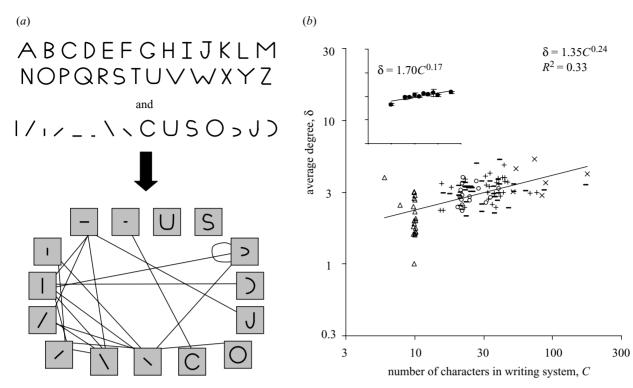


Figure 4. (a) Illustration of how a stroke-type network is built from the character repertoire and stroke type repertoire. Each stroke type is represented as a node in the network, and two stroke types are connected just in cases where those stroke types intersect in some character of the writing system; intuitively, stroke types sharing an edge in the network have the ability to 'interact'. When a stroke does not intersect other strokes of a character—like the dot of an 'i'—the stroke is deemed to intersect the physically nearest stroke. (b) Log–log plot of average stroke-type degree versus number of characters for 115 writing systems. Circles, abjad; plus symbols, abugida; minus symbols, alphabet; crosses, syllabary; and triangles, numerical. The linear regression line and equation are shown, along with correlation. *x*-axis values have been perturbed by $\pm 1\%$ to aid in their discrimination. The best-fit relationship is $\delta = 1.40C^{0.22}$ for invented systems (a set of independent data), and $\delta = 1.16C^{0.30}$ for non-invented systems. Inset: same plot, and same axes, but stroke-type degrees binned at 0.1 intervals along the log *C*-axis (standard error bars shown).

recognize objects, and many objects are built from object junctions of various kinds, which are themselves built out of contours of various types. Could it be that writing systems have been selected to have characters that can be recognized using the already-existing object recognition mechanisms? Intuitively, strokes are contour-like, and characters are object-junction-like, object junctions typically possessing approximately three intersecting contours (e.g. Clowes 1971; Huffman 1971; Chakravarty 1979). Might it be more than a coincidence that object junctions are well described as 'T', 'L', 'X', 'Ψ', 'K' and 'Y' junctions? A test of this hypothesis is the subject of ongoing research (Changizi *et al.* 2004).

We thank Dan Ryder, Wei Ji Ma, Patrick Wilken, Alan Hampton and two reviewers for their comments. We also thank Andrew Hsieh, Hao Ye and Qiong Zhang for independently parsing into stroke types a variety of writing systems.

REFERENCES

- Ager, S. 1998 Omniglot: a guide to writing systems. See http:// www.omniglot.com.
- Chakravarty, I. 1979 A generalized line and junction labeling scheme with applications to scene analysis. *IEEE Trans. Pattern Analysis Machine Intell.* 1, 202–205.
- Changizi, M. A. 2001 Universal scaling laws for hierarchical complexity in languages, organisms, behaviors and other combinatorial systems. *J. Theor. Biol.* 211, 277–295.

- Changizi, M. A. 2003*a* The relationship between number of muscles, behavioral repertoire, and encephalization in mammals. *J. Theor. Biol.* 220, 157–168.
- Changizi, M. A. 2003b The brain from 25 000 feet: high level explorations of brain complexity, perception, induction and vagueness. Dordrecht, The Netherlands: Kluwer.
- Changizi, M. A., McDannald, M. A. & Widders, D. 2002 Scaling of differentiation in networks: nervous systems, organisms, ant colonies, ecosystems, businesses, universities, cities, electronic circuits, and Legos. *J. Theor. Biol.* 218, 215–237.
- Changizi, M. A., Zhang, Q., Ye, H. & Shimojo, S. 2004 The structures of letters and symbols throughout human history are selected to match those found in objects in natural scenes. (Submitted.)
- Chen, Y. P., Allport, D. A. & Marshall, J. C. 1996 What are the functional orthographic units in Chinese word recognition: the stroke or the stroke pattern? *Q. J. Exp. Psychol. A Hum. Exp. Psychol.* **49**, 1024–1043.
- Clowes, M. B. 1971 On seeing things. Artificial Intell. 2, 79-116.
- Daniels, P. T. & Bright, B. 1996 *The world's writing systems*. New York: Oxford University Press.
- Dehaene, S. 1997 The number sense: how the mind creates mathematics. Oxford University Press.
- Huffman, D. A. 1971 Impossible objects as nonsense sentences. In *Machine intelligence*, vol. 6 (ed. B. Meltzer & D. Michie), pp. 295–323. New York: Elsevier.
- Ifrah, G. 1985 *From one to zero: a universal history of numbers.* New York: Viking Press.

- Pelli, D. G., Burns, C. W., Farell, B. & Moore, D. C. 2004 Identifying letters. *Vision Res.* (In the press.)
- Trick, L. M. & Pylyshyn, Z. W. 1994 Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychol. Rev.* 101, 80–102.
- Vogel, E. K., Woodman, G. F. & Luck, S. J. 2001 Storage of features, conjunctions, and objects in visual working memory. *J. Exp. Pscyhol. Hum. Percept. Perform.* 27, 92–114.
- Yeh, S. L. & Li, J. L. 2004 Sublexical processing in visual recognition of Chinese characters: evidence from repetition blindness for subcharacter components. *Brain and Language* 88, 47–53.
- Zhang, J. & Norman, D. A. 1995 The representation of numbers. *Cognition* 57, 271–295.

As this paper exceeds the maximum length normally permitted, the authors have agreed to contribute to production costs.