How to write around the world (And which ways are best)

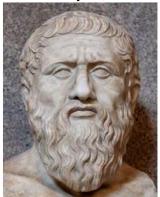
Jason Merchant

Professor, Department of Linguistics
Chair, Department of Slavic Languages and Literatures
Associate Dean for Languages
University of Chicago

Humanities Day, October 2015

Is there a right and wrong way to speak?

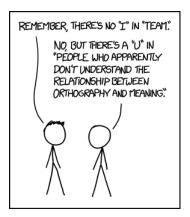
Plato's Cratylus, or On the correctness of names



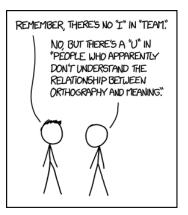
 Writing is a system to represent language (not 'thought') in visual form

 Writing is a system to represent language (not 'thought') in visual form

 Writing is a system to represent language (not 'thought') in visual form



 Writing is a system to represent language (not 'thought') in visual form



 The basic units are therefore *linguistic* units: words, morphemes, syllables, or phonemes

• Words and morphemes:

- Words and morphemes:
 - teachers: a word consisting of three morphemes

- Words and morphemes:
 - teachers: a word consisting of three morphemes
 - 2 teach-er-s: stem-suffix-suffix

- Words and morphemes:
 - teachers: a word consisting of three morphemes
 - 2 teach-er-s: stem-suffix-suffix
- Syllables and phonemes:

- Words and morphemes:
 - teachers: a word consisting of three morphemes
 - 2 teach-er-s: stem-suffix-suffix
- Syllables and phonemes:
 - This.sen.tence.does.n't.have.few.er.than.five.syl.la.bles.in.it.: Syllables are the basic unit of poetic meter (scansion), etc.

- Words and morphemes:
 - teachers: a word consisting of three morphemes
 - 2 teach-er-s: stem-suffix-suffix
- Syllables and phonemes:
 - This.sen.tence.does.n't.have.few.er.than.five.syl.la.bles.in.it.: Syllables are the basic unit of poetic meter (scansion), etc.
 - **Phonemes**: distinctive units of sound in a given language: consonants (*k*, *b*, *h*, ...); vowels (*a*, *e*, ...); diphthongs (*ai*, *au*, *oi*, ...); tones

Major types of writing systems

Туре	Symbols represent	Example
Logographic Syllabary Abjad	morpheme or word syllable C (V optional)	Chinese hànzì Japanese kana Hebrew, Arabic
Alphabet	consonants, vowels	Greek, Latin, Cyrillic

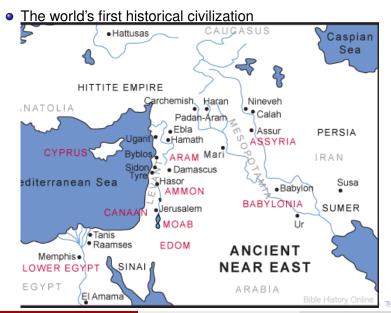
Alphabets and abjads are phonemic systems

Major types of writing systems

Туре	Symbols represent	Example	
Logographic	morpheme or word	Chinese <i>hànzì</i>	
Syllabary	syllable	Japanese <i>kana</i>	
Abjad	C (V optional)	Hebrew, Arabic	
Alphabet	consonants, vowels	Greek, Latin, Cyrillic	

- Alphabets and abjads are phonemic systems
- Phonemic systems and syllabaries are phonographic systems

Sumerian: The language of Sumer, ca. 3000-2000 BC



 'Cuneiform' means 'wedge-shaped', from the shapes the stylus made in clay



 Began as pictographs (stylized representations of objects or properties thereof or relations between)





 Began as pictographs (stylized representations of objects or properties thereof or relations between)





Became logographic (one symbol = one morpheme/word)

 Began as pictographs (stylized representations of objects or properties thereof or relations between)





- Became logographic (one symbol = one morpheme/word)
- First place we see evidence of phonographic writing: the use of symbols to represents sound (syllables or segments) rather than just concepts or ideas (3000-2700 BCE)

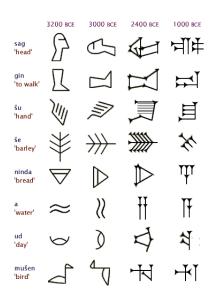
 Began as pictographs (stylized representations of objects or properties thereof or relations between)





- Became **logographic** (one symbol = one morpheme/word)
- First place we see evidence of phonographic writing: the use of symbols to represents sound (syllables or segments) rather than just concepts or ideas (3000-2700 BCE)
- Made possible by the use of the rebus principle: a symbol could be used for any word that was pronounced like the word whose meaning it originally represented.

From logograms to cuneiform



Meanwhile, the Egyptians were up to something similar
 Rosetta Stone: Jean-François Champollion





Meanwhile, the Egyptians were up to something similar
 Rosetta Stone: Jean-François Champollion





• Hieroglyphics mixed logographs....













Meanwhile, the Egyptians were up to something similar
 Rosetta Stone: Jean-François Champollion





• **Hieroglyphics** mixed logographs....



Determinatives...

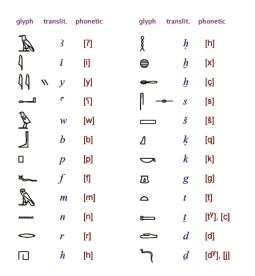


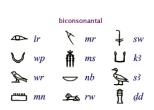


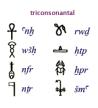




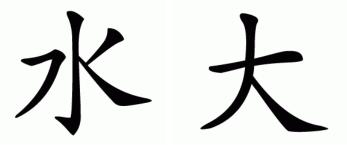
• And, eventually, some phonographs:







 Originally pictographs, developed into logographs: what do the following mean?



 Originally pictographs, developed into logographs: what do the following mean?



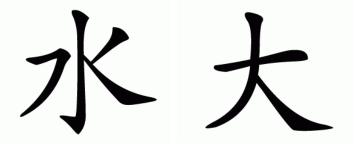
In Pinyin transcription: shui³, da⁴

 Originally pictographs, developed into logographs: what do the following mean?



- In Pinyin transcription: shui³, da⁴
- 'water' 'big'

 Originally pictographs, developed into logographs: what do the following mean?



- In Pinyin transcription: shui³, da⁴
- 'water' 'big'
- Today most characters consist of two parts: a semantic radical (bùshŏu) and phonetic determinative

Maya glyphs: Writing of the Mayan empire, 2000BC-900AD



• Palenque:

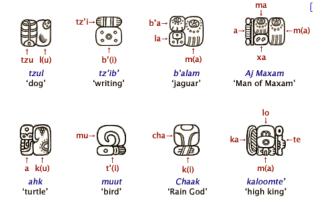




Partly logographic:



Partly syllabic:



Sometimes both



Japanese katakana: Mostly a syllabary

Katakana syllabograms

	Monographs (gojūon)						
	а	i	u	e	0		
Ø	ア	1	ウ	エ	オ		
	a [a]	i [i]	u [u]	e [e]	o [o]		
K	カ	+	ク	ケ	コ		
	ka [ka]	ki [ki]	ku [ku]	ke [ke]	ko [ko]		
s	サ	シ	ス	セ	ソ		
	sa [sa]	shi [si]	su [su]	se [se]	so [so]		
T	タ	チ	ツ	テ	۲		
	ta [ta]	chi (tci)	tsu [tsu]	te [te]	to [to]		
N	ナ	=	ヌ	ネ	1		
	na [na]	ni [nʲi]	nu [nu]	ne [ne]	no [no]		

Cherokee: 86 symbols (syllabographs)

Invented from scratch in the early 1800s by Sequoyah



Cherokee: 86 symbols (syllabographs)

Invented from scratch in the early 1800s by Sequoyah



Almost a true syllabary (find the exception):
 (The 'v' in the last column stands for a front mid unrounded nasalized vowel)

Cherokee: 86 symbols (syllabographs)

D a	R e	T i	δ .	O u	ì v
S ga O ka	I ⁰ ge	$oldsymbol{y}$ gi	A go	$oldsymbol{J}$ gu	${f E}$ gv
${ m o}{m V}$ ha	\mathbf{P} he	A hi	I ho	Γ hu	Q hv
W la	δ le	P ii	G lo	M lu	a 1v
o <mark>t∽</mark> ma	Ol me	\mathbf{H} mi	5 mo	y ™u	
Θ na t _π hna C nah	∫ ne	\mathbf{h} ni	Z no	q nu	Op un
T qua	${\mathfrak O}$ dne	${f \Omega}$ dai	∜ quo	(C) dnn	8 quv
oÐ ∘ H ∞	4 se	${f P}$ si	P so	$oldsymbol{g}$ su	${f R}$ sv
$oldsymbol{\zeta}$ da $oldsymbol{W}$ ta	S de b te	$oldsymbol{J}$ di $oldsymbol{J}$ ti	Λ do	\mathbf{S} du	$Q_{\mathbf{o}} \text{q} \wedge$
& dla Γ tla	L tle	C tii	V tlo	1 tlu	P tlv
G tsa	${f V}$ tse	${f h}$ tsi	K tso	$oldsymbol{d}$ tsu	C" tsv
G wa	Q9 we	(1) wi	ow CO	9 wu	6 w
Ø ya	β уе	Љ yi	h yo	G yu	В уч

Syllabaries and phonemic inventories

 Suspect anything about the phonological properties of these languages?

Syllabaries and phonemic inventories

- Suspect anything about the phonological properties of these languages?
- Hint: How many vowels do they have? How many consonants? How do these combine?

Syllabaries and phonemic inventories

- Suspect anything about the phonological properties of these languages?
- Hint: How many vowels do they have? How many consonants? How do these combine?
- A syllabary would be a nightmare for a language like English...
 how many distinct possible syllables does English have?

The origins of writing in Greece

• Earliest deciphered Greek: Linear B (ca. 1400-1200 BC)



 Developed by Semitic peoples somewhere in Phoenicia (now Lebanon) 1000 BCE?

- Developed by Semitic peoples somewhere in Phoenicia (now Lebanon) 1000 BCE?
- First major phonemic script

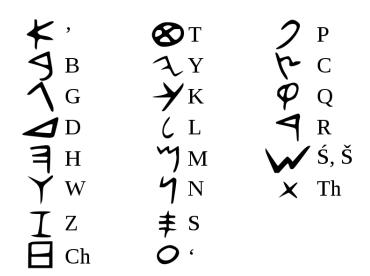
- Developed by Semitic peoples somewhere in Phoenicia (now Lebanon) 1000 BCE?
- First major phonemic script
- However, it was still an abjad: no vowels

- Developed by Semitic peoples somewhere in Phoenicia (now Lebanon) 1000 BCE?
- First major phonemic script
- However, it was still an abjad: no vowels
- Descended from pictographic Proto-Sinaitic via Proto-Canaanite

Phoenician

Name	Proto-Canaanite	Phoenician	Gloss
'aleph	J	4	'ox'
taw	+	+	'mark, signature'
mem	~~~	4	'water'

Phoenician



Hebrew

×	'alep	,	ל	lamed	1
コ	bet	b	מם	mem	m
٦	gimel	g	בך	nun	n
7	dalet	d	D	samek	s
7	he	h	Y	ayin	•
٦	waw	w	פף	pe	p
7	zayin	z	צץ	tsade	ş
Π	het	þ	7	qop	q
U	tet	ţ	٦	reš	r
7	yod	у	\mathtt{w}	šin	š
כך	kap	k	ת	taw	t

• S th strctr f nglsh n mr r lss clr wtht th vwl lttrs?

- S th strctr f nglsh n mr r lss clr wtht th vwl lttrs?
- Wht's wrs: nglsh wrttn wth n vwls, r...

- S th strctr f nglsh n mr r lss clr wtht th vwl lttrs?
- Wht's wrs: nglsh wrttn wth n vwls, r...
- Englesh wretten en sech e wee thet evere vewel es repleced be e sengle sembel?

- S th strctr f nglsh n mr r lss clr wtht th vwl lttrs?
- Wht's wrs: nglsh wrttn wth n vwls, r...
- Englesh wretten en sech e wee thet evere vewel es repleced be e sengle sembel?
- Duus thu fruquuncu uf thu sumbul muttur?

- S th strctr f nglsh n mr r lss clr wtht th vwl lttrs?
- Wht's wrs: nglsh wrttn wth n vwls, r...
- Englesh wretten en sech e wee thet evere vewel es repleced be e sengle sembel?
- Duus thu fruquuncu uf thu sumbul muttur?
- What abadt af Q ward to pack a mard arbatrard sambal?

- S th strctr f nglsh n mr r lss clr wtht th vwl lttrs?
- Wht's wrs: nglsh wrttn wth n vwls, r...
- Englesh wretten en sech e wee thet evere vewel es repleced be e sengle sembel?
- Duus thu fruquuncu uf thu sumbul muttur?
- What abadt af Q ward to pack a mard arbatrard sambal?
- I e uue o ei ay oe o e ea iou e oe ee?

A useful development

Phoenician was an abjad...

A useful development

- Phoenician was an abjad...
- But the Greeks borrowed the symbols and made the first true alphabet (<alpha + bet(a))

A useful development

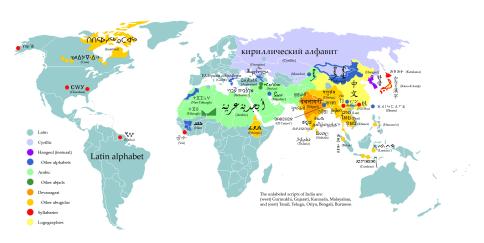
 The Greeks re-appropriated certain symbols to serve their phonemic needs.

Ph	noeniciar	า		Greek	
symbol	name	value	symbol	name	value
>	'ālep	[7]	$\mathbf{A}\alpha$	alpha	[a]
7	he	[h]	$E\varepsilon$	epsilon	[e]
	yodh	[j]	1ι	iota	[i]
F	heth	[ħ]	$H\eta$	eta	[13]
0	'ayin	[ና]	O <i>o</i>	omicron	[o]
Y	waw	[w]	Yv	upsilon	[y] ([i])

The Greek alphabet

Aα	Alpha	Νv	Nu
Ββ	Beta	Ξξ	Xi
Γγ	Gamma	Oo	Omicron
$\Delta \delta$	Delta	$\Pi\pi$	Pi
Eε	Epsilon	PQ	Rho
$\mathbf{Z}\zeta$	Zeta	Σσς	Sigma
Ηη	Eta	$T\tau$	Tau
Θθ	Theta	Yυ	Upsilon
Ιι	lota	$\Phi \phi$	Phi
Kκ	Kappa	Xχ	Chi
Λλ	Lambda	$\Psi \psi$	Psi
Mu	Mu	Ω_{0}	Omena

Alphabets from the Greek: Latin, Cyrillic, etc...



The Russian (Cyrillic) alphabet

Аа Да Бб Б В В В Гг Гг Дд Дд Ее Ее Ёё Её Жж Жж Зз Зз Ии Ии Ййй Кк Кл Лл da Mm Mu HH Hn Oo Oo Пп \mathcal{I}_n Рр \mathcal{P}_n Сс \mathcal{C}_c Тт \mathcal{I}_m Уу Уу ФФ Фр Хх Хх Цц Цц Чч Уч ШшШш ЩщШш Ъъъ Ыы а Ьь 6 ЭэЭэ Юю Юю Яя Ях,

Our own alphabet: the Roman or Latin

ABCDEFGHIJKLMNOPQRSTUVWXYZ

• 26 letters, well suited to writing Latin: how about English?

	Bilabial	Labio- dental	Dental	Alveolar	Post- alveolar	Palatal	Velar	Glottal
Plosive	рb			t d			k g	
Affricate					t∫ dʒ			
Fricative		f v	θð	s z	J 3			h
Nasal	m			n			ŋ	
Lateral				1				
Approximant					I	j	(w) w	

• Cf. tin, thin, thing, this, thick, tick, sick: 3 sounds in each

	Bilabial	Labio- dental	Dental	Alveolar	Post- alveolar	Palatal	Velar	Glottal
Plosive	рb			t d			k g	
Affricate					t∫ dʒ			
Fricative		f v	θð	s z	J 3			h
Nasal	m			n			ŋ	
Lateral				1				
Approximant					ı	j	(w) w	

- Cf. tin, thin, thing, this, thick, tick, sick: 3 sounds in each
- breath, breathe, meth, seethe, bath, bathe

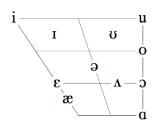
	Bilabial	Labio- dental	Dental	Alveolar	Post- alveolar	Palatal	Velar	Glottal
Plosive	рb			t d			k g	
Affricate					t∫ dʒ			
Fricative		f v	θð	s z	J 3			h
Nasal	m			n			ŋ	
Lateral				- 1				
Approximant					ı	j	(w) w	

- Cf. tin, thin, thing, this, thick, tick, sick: 3 sounds in each
- breath, breathe, meth, seethe, bath, bathe
- <u>th</u> is a digraph: two letters used to represent one sound

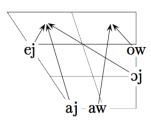
	Bilabial	Labio- dental	Dental	Alveolar	Post- alveolar	Palatal	Velar	Glottal
Plosive	рb			t d			k g	
Affricate					t∫ dʒ			
Fricative		fv	θð	s z	J 3			h
Nasal	m			n			ŋ	
Lateral				1				
Approximant					I	j	(w) w	

- Cf. tin, thin, thing, this, thick, tick, sick: 3 sounds in each
- brea<u>th</u>, brea<u>th</u>e, me<u>th</u>, see<u>th</u>e, ba<u>th</u>, ba<u>th</u>e
- <u>th</u> is a digraph: two letters used to represent one sound
- Old English had letters for these: þ ('thorn') and ð ('eth'): ðis <this> vs. þin <thin>

Phonemes of Midwestern American English: 15 vowels and diphthongs



beat, bit, bet, bat, boot, put, but, bought, (ro)bot, a(bout)



 $b\underline{ai}t,\,b\underline{i}te,\,b\underline{ou}t,\,B\underline{oy}d,\,b\underline{oa}t$

• gh as in enough

- gh as in enough
- o as in women

- gh as in enough
- o as in women
- ti as in nation

- gh as in enough
- o as in women
- ti as in nation
- = fish!

Is texting destroying the English language?

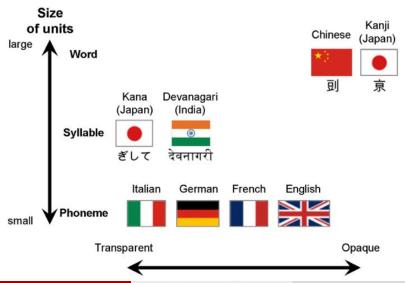
Fewer than 10% of words are abbreviated in texts

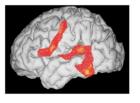
- Fewer than 10% of words are abbreviated in texts
- Abbreviations have always been with us

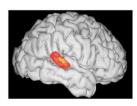
- Fewer than 10% of words are abbreviated in texts
- Abbreviations have always been with us
- To abbreviate presupposes some spelling

- Fewer than 10% of words are abbreviated in texts
- Abbreviations have always been with us
- To abbreviate presupposes some spelling
- Texting is supplementing, not supplanting, speech; it is not replacing any equivalent form of written communication

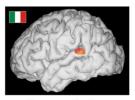
- Fewer than 10% of words are abbreviated in texts
- Abbreviations have always been with us
- To abbreviate presupposes some spelling
- Texting is supplementing, not supplanting, speech; it is not replacing any equivalent form of written communication
- Autocomplete is already ending abbreviations







Network common to reading in Italian and in English



Stronger activation in Italian



Stronger activation in English

Figure 2.20. The transparency of the spelling system influences the organization of the reader's brain. Writing systems differ in the size of the units they denote (phonemes, syllables, or whole words) and in their degree of transparency (the regularity of the relation between symbols and speech sounds). Italian spelling is very regular, while English spelling bristles with irregularities and exception words. When contrasting brain activations in Italian and English, small modulations are seen within an overall shared network. Italian causes stronger activation in auditory areas of the superior temporal lobe, while English puts greater emphasis on the

'Whole word' reading strategies are inferior to phonemic training



Figure 5.2. In spite of its inefficacy, the whole-language method continues to inspire educators throughout the world. The tests shown here are still used in some classrooms during the first few weeks of reading instruction. Although the child cannot yet sound out letters, he is asked to pair a whole word with the corresponding image (top). He of course makes gross errors such as calling a cat "pig," suggesting that he is unaware of the alphabetic principle. The child is also taught to attend to the overall contours of ascending and descending letters. Note that even the teacher errs on the word "boat." These exercises are in no way related to how our brain recognizes words.



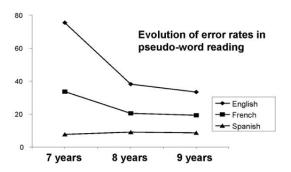


Figure 5.3. European languages are not equal in the face of reading acquisition. Error rates in reading familiar words were measured in fifteen European countries after one year of schooling (top, data from Seymour et al., 2003, table 5). Finnish, German, Greek, Austrian and Italian, whose spelling is transparent, were already read accurately. At the other end of the scale, English is very opaque and children could only read one out of three words. These inequalities were perpetuated in subsequent years, particularly for pseudo-words such as "balist" or "chifling" that can only be deciphered with grapheme-phoneme decoding (bottom, data from Goswami et al., 1998, table 8). It takes one or two additional years of schooling before an English child reaches the reading level of a French child.

So English writing (and spelling) can be tough to learn... but at least it's not Irish:

Rachaidh mé 'I WILL go.'

So English writing (and spelling) can be tough to learn... but at least it's not Irish:

- Rachaidh mé 'I WILL go.'
- Coinneochaidh mé 'I WILL keep.'

So English writing (and spelling) can be tough to learn... but at least it's not Irish:

- Rachaidh mé 'I WILL go.'
- Coinneochaidh mé 'I WILL keep.'
- ach tiocfaidh sé 'but he WILL come'

So English writing (and spelling) can be tough to learn... but at least it's not Irish:

Rachaidh mé 'I will go.' 'rahə 'meː

So English writing (and spelling) can be tough to learn... but at least it's not Irish:

- Rachaidh mé 'I WILL go.'
- Coinneochaidh mé

'I WILL keep.'

'rahə 'me:

'kin^jahə 'meː

So English writing (and spelling) can be tough to learn... but at least it's not Irish:

- Rachaidh mé 'I WILL go.'
- Coinneochaidh mé 'I WILL keep.'
- ach tiocfaidh sé 'but he WILL come'

'rahə 'meː

ˈkiɴ^jahə ˈmeː

't^jukə 'ʃeː

Thank you and...



Enjoy your visit!

The Disk of Phaistos: 1800-1600 BC, found on Crete



Logographic, syllabary, abjad, alphabet?

The Disk of Phaistos: 1800-1600 BC, found on Crete



- Logographic, syllabary, abjad, alphabet?
- 241 stamped tokens, 45 unique signs

The Disk of Phaistos: 1800-1600 BC, found on Crete



- Logographic, syllabary, abjad, alphabet?
- 241 stamped tokens, 45 unique signs
- Most likely a syllabary