

PSYCH-UH 2218: Language Science

Class 11: Writing Systems

Prof. Jon Sprouse Psychology

Writing Systems

A writing system is a method of **representing** a language visually on a physical medium (paper, papyrus, clay, stone, etc).

Three linguistic questions we can ask about writing systems:

1. How many times has writing been invented?

(And what does this tell us about the relationship between language and writing?)

2. How many different kinds of writing systems are there?

(And how do they relate the linguistic representations that we have been studying in this course?)

3. How did writing systems spread throughout the world?

(And how do they interact with the grammars of the languages that adopt them?)

How many times has writing been invented?

Not the invention of a specific system, the invention of the idea

When we ask how many times writing has been invented, we don't mean a specific instantiation of a writing system.

People invent writing systems all of the time. Here are two recently invented writing systems (for constructed languages):

Klingon

E	*(Ŋ	১	Γ	I	ৎ	९	ת	L
а	b	ch	D	е	gh	Н	I.	j	Т
[a]	[b]	[ʧ]	[d]	[٤]	[¥]	[x]	[1]	[අ]	[1]
9	6	3	3	う	\$	5	7	X	M
m	n	ng	ο	р	q	Q	r	S	t
[m]	[n]	[ŋ]	[0]	[p ^h]	[q ^h]	$[\widehat{q\chi}]$	[r]	[ş]	[t ^h]
							Punctu	tation	
Y	4	Ж	て	र	S		•	▲	ø
e th	L u) (۲ «	Q y	<u>۲</u>		▼ pause	▲ period	ø
१ th [t]	u [u])(v [v]	С [w]	Q y [j]	, [3]		▼ pause	▲ period	ø
لا th [tł]	u [u]) v [v]	C w [w]	Q у [ј]) [7]		▼ pause	▲ period	ø
th [tł] Numera	u [u] his)(v [v] (ר יי ייי ייי	ک	ہ [ی] ک	ب ر	▼ pause	▲ period	
th [tł] Numera pagh	u [u] his vva')(v [v] (cha'	C w [w] ¢ wej	Р у []] Зор	2 [?] & vagh) ↑ jav	▼ pause * (Soch	period	♪ Hut

Elvish (Quenya) Tengwar script

p	p	Ч	Ч	ps	pp	ccj	æ
picel	pym	ç ĝ	ษ์ว์	ົຳກຳ	íphy	îcêj	íщ
tinco	parma	calma	quesse	ando	umbar	anga	ungwe
t	р	С	qu	nd	mb	ng	ngw
[t]	[p]	[k]	$[k^w]$	[nd]	[mb]	[ŋg]	[ŋ g ^w]
b	b	cl	ਖ	bo	b	ccl	ਕ
bjć	bymm	ĉlym	đĵ	îĥ	î	îcĉĺ	îtt
súle	formen	harma	hwesta	anto	ampa	anca	unque
s	f	ch	hw	nt	mp	nc	nqu
[s]	[f]	[X]	[x ^w]	[nt]	[mp]	[ŋk]	[ŋ k ^w]
177	m	cci	ω	n	n	С	ದ
ာာက်ဘာ	ŵŢŖ	ŵ£	ய்க ும்	îń	ΰĈ,	îŵ	ن څ
númen	malta	ngoldo	ngwalme	ore	vala	anna	vilya
n	m	n	nw	r	v	У	v/w
[n]	[m]	[n]	$[n^w]$	[r]	[v]	[j]	[v/w]
y	Þ	\overline{C}	5	С	2	ģ	3
<u> </u>	îŵ	ĉώ	îŝ	ή'n	ήm	îś	îś
C	c	- /	-		က်င်္ခဗက်	12	າກ ໌ ຊູ່yກື
romen	arda	lambe	alda	silme	silme	áze	áze
					nuquerna		nuquerna
r	rd	1	ld	s	s	z	z
[r]	[rd]	[1]	[ld]	[s]	[s]	[z]	[z]
λ	d	$\boldsymbol{\lambda}$	0	c	ſ	1	1
λήμω	ďځ	ૌmp	ĵó	ćģ	ΧĜ	ŕτý	ĴÊ
	ؿؿؗۘ؇ۣۺۯ						
hyamen	hwesta sindarinwa	yanta	úre	osse	halla	telco	ára
hy	hw	i	u		h	short vowel	long vowel
[j/ç]	[w/h ^w]	[i]	[u]		[h]	carrier	carrier

There are 3 undisputed instances of the invention of the **idea** of writing



And 2 instances that are debated (because of trade with Mesopotamia)



What about all of the other people, and all of the other times that people existed?



Do we think that people did not have language before writing was invented? Do we think that people did not have language before the idea of writing was spread to their culture? No, of course not. **Language predates writing.** This shows us that **language is independent** of writing, but **writing is dependent on language**.

The dissociation between language and writing

The idea of literacy rates is another important piece of evidence that language is distinct from writing. Literacy rates are not 100%. Because reading must be taught explicitly to humans. But have you ever heard anyone talk about "speaking rates"? No, that feels absurd. Because it is always 100% (in the absence of disease, injury, disorder, or abuse).



Source: WDI, CIA World Factbook, & other sources

Note: Specific definitions and measurement methodologies vary across countries and time. See the 'Sources'-tab for more details.

OurWorldInData.org/literacy • CC BY

How many different **types** of writing systems are there?

(This website has lots of information and graphics of different writing systems: <u>https://omniglot.com/</u>)

Let's think about linguistic levels

If we are going to use symbols to represent linguistic units, we can ask which linguistic level the symbols represent.

Here I am using a hashmark (#) to indicate a unit in each level that a written symbol could represent:

"words" /raitıŋ sıstəmz ar kampləketəd θıŋz/ #.# #.# # #.#.# #.# morphemes /rait.in sistem.z ar kampleket.ed 0in.z/ #.# #.# # #.#.#.# # syllables /rai.tıŋ sıs.təmz ar kam.plə.ke.təd 0ıŋz/ #.#.#.#.#.#.#.#.#.#.#.# #.# #.# #.#.#.#.#.#.#.#.#.#.#.#.#.#.#.# segments /r.ai.t.ı.ŋ s.ı.s.t.ə.m.z a.r k.a.m.p.l.ə.k.e.t.ə.d 0.ı.ŋ.z/ (phonemes)

Writing systems are seldom perfectly uniform

This is important to keep in mind. As we look at different writing systems, they will typically be dominated by a specific approach. But they may not be only that one approach. They may mix other approaches. Most writing systems are not perfectly uniform (other than the IPA, which was specifically created by scientists to be uniform).

/raitıŋ sıstəmz ar kampləketəd θıŋz/ "words" #.# #.# # #.#.# #.# morphemes /rait.in sistem.z ar kampleket.ed 0in.z/ #.# #.# # #.#.#.# # syllables /rai.tıŋ sıs.təmz ar kam.plə.ke.təd 0ıŋz/ #.#.#.#.#.#.#.#.#.#.#.# #.# #.# #.#.#.#.#.#.#.#.#.#.#.#.#.#.# phonemes /r.ai.t.ı.ŋ s.ı.s.t.ə.m.z a.r k.a.m.p.l.ə.k.e.t.ə.d 0.ı.ŋ.z/

Abjads: segment level, consonants

An abjad is a writing system that only represents consonants. It is named after the first four letters of the Arabic writing system (abjd) because Arabic and Hebrew are two of the most famous abjads.

Here is the Arabic abjad.

Remember that, in practice, no writing system is completely uniform. Many abjads, including Arabic, do include diacritics to indicate vowels in some texts, particularly for religious texts and other texts where ambiguity would be a problem.

5	د	د	Ċ	ζ	ご	ٹ	ت	ب	١
راء	ذال	دال	خاء	حاء	جيم	ثاء	تاء	باء	ألف
rā'	dāl	dāl	ḫā'	<u></u> hā'	ģīm	ţā'	tā'	bā'	'alif
r	đ	d	b	Ļ.	ğ	t	t	b	'(a)
[r~r]	[ð]	[d]	[x~\chi]	[ħ]	[ዓ]	[θ]	[t]	[b]	[?]
ف	Ż	ع	ظ	ط	ض	ص	ش	س	ز
فاء	غين	عين	ظاء	طاء	ضاد	صاد	شين	سين	ز اي
fā'	ġayn	ʻayn	<u></u> zā'	ţā'	ḍād	şād	šīn	sīn	zāy
f	ġ	¢	Ż	ţ	ġ	ş	š	s	Z
[f]	[Х~r]	[?]	[ð [°]]	[t [°]]	[d ^s]	[s [°]]	[ʃ]	[s]	[z]
	ç	ي	و	٥	ن	م	J	ای	ق
	همزة	ياء	واو	هاء	نون	ميم	لام	کاف	قاف
	hamza	yā'	wāw	hā'	nūn	mīm	lām	kāf	qāf
		у	w	h	n	m	I.	k	q
		[j]	[w]	[h]	[n]	[m]	[l~t]	[k]	[q]

Abjads and templatic morphology

This is the Hebrew Abjad.

Recall that both Arabic and Hebrew are classic examples of languages that use **templatic morphology**.

In templatic morphology, the consonants form the roots, and the vowel tier adds something like inflectional morphology.

It is probably not a coincidence that languages with templatic morphology tend to use abjads as their writing systems. One idea is that the root needs to be represented because it is not predictable, but the inflectional morphology given by the vowels can be predicted from the context of the sentence, so it does not need to be explicitly represented.

n	7	٦	n	٦	ړ	ב/ב	א
חֵית	זין	ŋ	הַא	דָּלֶת	גִּימֶל	בֵּית/בֵית	אָלֶף
het	zayin	vav	he	daled	gimel	bet	'alef
ħ	Z	v	h	d	g	b/v	,
[x~\chi]	[z]	[v/w]	[h~?]	[d]	[9]	[b/v]	[?]
8	7	6	5	4	3	2	1
ע	D	נ	n	ל	<u>)</u>	>	υ
עין	סָמֶךְ	נוּך	מֵם	לָמֶד	ڊِp/ڊ	יוֹד	טַית
'ayin	samekh	nun	mem	lamed	kaf/khaf	yod	tet
•	s	n	m	I.	k/kh	У	t
[?]	[s]	[n]	[m]	[1]	[k/x~ ₂]	[j]	[t]
70	60	50	40	30	20	10	9
		ת	ש/ש	٦	マ	צ	ฏ∕ฏ
		ໆກִ∕າກຼ	שִׁין/ש∶ין	רַישׁ	קוֹף	צָדִי/ צדיק	פַא/פה
		taw	shin	resh	qof	zadi	pe/fe
		t	s/sh	r	k	Ž	p/f
		[t]	[s/ʃ]	[Å∽R]	[k]	[ts]	[p/f]
		400	300	200	100	90	80
						Final co	nsonants
		۲	ງ	٦	Þ	T	Ŧ
		צָדִי סוֹפִית	פַא סוֹפִית	נוּן סוֹפִית	מַם סוֹפִית	כָף סוֹפִית	כָף סוֹפִית
		zadi sofit	pe sofit	nun sofit	mem sofit	khaf sofit	kaf sofit
		final <u>z</u> adi	final pe	final nun	final mem	final khaf	final kaf
		Co	ommon ortho	graphy addi	tions (mainly	for foreign b	oorrowings)
וי/וו	תי	צי	רי/עי	'n	15	'7	גי
w	th	č	gh	kh	ž	dh	ğ
[w]	[θ]	[ʧ]	[R]	[x]	[3]	[ð]	[ʤ]

Alphabets: segment level, C & V

Alphabets represent every segment - both consonants and vowels. The segments are typically phonemes (not allophones), though this can vary.

We are all familiar with the Latin alphabet, so I've put this cool graphic about its evolution here.

We will talk about the history of writing systems later today.



Syllabaries: syllable level

In a syllabary, each symbol represents a syllable. Obviously, what counts as a syllable is dictated by the phonology of the language: V, CV, CVC, etc.

This is the syllabary for Cherokee, a Native American language.

This syllabary was developed in the early 1800s.

D	а				R	е		Т	i		൴	o	O°	u	i	v
S	ga	9	ka		þ	ge		У	gi		А	go	J	gu	E	gv
Փ	ha				P	he		A	hi		ŀ	ho	Γ	hu	ŀ	hv
W	la				ഗ	le		Ր	li		G	lo	М	lu	ጓ	lv
$\mathfrak{I}_{\mathbf{r}}$	ma				Ю	me		Η	mi		5	mo	Y	mu		
θ	na	ե	hna	G $^{\mbox{\tiny nah}}$	\mathbb{V}	ne		h	ni		Ζ	no	Ð	nu	Or	nv
Τ	qua				ω	que		թ	qui		vr	quo	ശ	quu	3	quv
Ĥ	sa	$\mathbf{G}_{\mathbf{O}}$	s		4	se		Р	si		ተ	so	୪	su	R	sv
Ն	da	W	ta		Ց	de	b te	Л	di	Лü	V	do	S	du	ന	dv
ൽ	dla	Ĺ	tla		L	tle		С	tli		મ	tlo	P	tlu	Р	tlv
C	tsa				\mathcal{V}	tse		h	tsi		Κ	tso	Ъ	tsu	Cĩ	tsv
G	wa				QI	we		0	wi		v	wo	I	wu	6	wv
ධ	ya				B	ye		$\sqrt{2}$	yi		ĥ	уо	Gĩ	yu	В	yv

Syllabaries: syllable level

This is hiragana, one of the two syllabaries used to write Japanese. It was developed from Chinese characters, called kanji in Japanese. This chart shows a romanized value on the left, the hiragana symbol in the middle, and the kanji symbol that it was likely derived from on the right.

平化	反名	(ひ	らた	がな) hii	ragar	na							
a	あ	安	i	い	以	u	う	宇	e	え	衣	0	お	於
ka	か	加	ki	き	幾	ku	<	久	ke	け	計	ko	٦	5
sa	さ	左	shi	し	之	su	す	寸	se	せ	世	so	そ	曽
ta	た	太	chi	ち	知	tsu	つ	Л	te	τ	天	to	٤	止
na	な	奈	ni	に	仁	nu	ぬ	奴	ne	ね	袮	no	の	乃
ha	は	波	hi	ひ	比	fu	ふ	不	he	\sim	部	ho	ほ	保
ma	ま	末	mi	み	美	mu	む	武	me	め	女	mo	も	毛
ya	や	也				yu	ゆ	由				уo	よ	与
ra	6	良	ri	IJ	利	ru	る	留	re	れ	礼	ro	ろ	呂
wa	わ	和	wi	ゐ	為				we	ゑ	恵	wo	を	遠

Syllabaries and syllable inventory

Syllabaries require one symbol for each syllable. This means that they will tend to have more symbols than alphabets. This also means that they will tend to work best for languages that have smaller syllable inventories.

We see this with Japanese, which we have already seen has a relatively small syllable inventor (mostly V or CV).

For languages with larger syllable inventories, like many Indo-European languages, alphabets will be a more economical system otherwise, the number of symbols will grow quite large. 平仮名(ひらがな) hiragana

a	あ	安	i	い	以	u	う	宇	e	え	衣	0	お	於
ka	か	加	ki	き	幾	ku	<	久	ke	け	計	ko	Ľ	5
sa	さ	左	shi	し	Ż	su	す	寸	se	せ	世	so	そ	曽
ta	た	太	chi	ち	知	tsu	つ	Л	te	τ	天	to	٤	ıЕ
na	な	奈	ni	に	仁	nu	ぬ	奴	ne	ね	袮	no	の	乃
ha	は	波	hi	ひ	比	fu	ふ	不	he	\sim	部	ho	ほ	保
ma	ま	末	mi	み	美	mu	む	武	me	め	女	mo	も	毛
ya	や	也				yu	ゆ	由				уo	ደ	与
ra	6	良	ri	IJ	利	ru	る	留	re	れ	礼	ro	ろ	呂
wa	わ	和	wi	ゐ	為				we	ゑ	恵	wo	を	遠
												n	h	无

Abugidas: between segments and syllables

Abugidas represent both consonants and vowels, but the system is complicated. (The name comes from the first 4 letters of the Ethiopian system: abgd)

The primary symbols each represent a consonant plus an "inherent" vowel. In that sense, the symbols represent syllables.

The non-default vowels are then indicated by diacritics that are added to the primary symbols. In that sense, the symbols are alphabetic.



The vowels can also be indicated with their own symbol when they appear at the beginning of a word (not following a consonant), like an alphabet.

Because abugidas have properties of both alphabets and syllabaries, they are sometimes called syllabic alphabets.

Abugidas: between segments and syllables

An example of an abugida is the Devanāgarī script used to write Hindi and many other languages in south and southeast asia. Here are the vowels:

Primary vowels													
_		Short				Lo	ing		[Diphtł	nongs		
	Initi	al	Diac	ritic	Initia	al	Diac	ritic	Initi	al	Diac	ritic	
Unrounded low central	अ	а	प	ра	आ	ā	पा	pā					
Unrounded high front	इ	i	पि	pi	ई	ī	पी	pī					
Rounded high back	उ	u	पु	pu	ऊ	ū	पू	рū					
Syllabic variants	ॠ	ŗ	पृ	pŗ	ॠ	ŗ	Ţ	pŗ					
	ऌ	ļ	पू	рĮ	ॡ	Î	पू	pĮ					
Secondary vowels													
Unrounded front					ए	е	पे	pe	ऐ	ai	पै	pai	
Rounded back					ओ	0	पो	ро	औ	au	पौ	pau	

Abugidas: between segments and syllables

An example of an abugida is the Devanāgarī script used to write Hindi and many other languages in south and south east asia. Here are the consonants:



Logographic systems...

The name "logographic" means that that each symbol (graph) represents a word or idea (logo).

This idea is very easy to grasp, and very seductive. Here are examples from Egyptian hieroglyphs (top) and Chinese characters (bottom) that illustrate this. There is even a pictographic relationship the symbol looks like the concept.



But, in practice, what we call logographic systems are often much more complicated.

The symbols for words tend to be **compounds** - made up of multiple symbols.

And when we look inside the compounds, we tend to see both **semantic** and **phonetic** components to the compounds. 媽

ma - mother

Logographic systems and semantic-phonetic compounds

The precise interaction of semantic and phonetic components varies by writing system (Egyptian hieroglyphs work differently from Mayan glyphs and from Chinese characters). But we can illustrate the general idea with a classic example from Chinese characters.





This is the symbol for horse. It is <u>not</u> a compound. It has the meaning 'horse' (which you can kind of see in the shape) and the pronunciation 'ma'.



This is the symbol for cart. It <u>is a compound symbol</u>. Its phonetic component is 'jia', but it contains the horse symbol as a <u>semantic component</u>.



This is the symbol for mother. It is a <u>compound symbol</u>. Its <u>phonetic component</u> is 'ma', which is signified by the symbol for horse. The other symbol is 'woman', which contributes to the semantic component.

How did writing systems spread throughout the world?

Sumerian Cuneiform ~3300 BCE

The Sumerians moved into Mesopotamia (Iraq/Syria) around 4000 BCE. They displaced a different people that we know very little about.

They built vast cities and created complex economies!

For example, at one point they started using clay tokens to represent goods that were being bought/sold.

They would put those clay tokens inside of clay balls that acted as sealed envelopes. If the contract came under dispute, they could break open the ball to see what was originally agreed to.

Eventually, thy started imprinting the tokens on the outside of the envelope so everyone could see what was agreed to without actually breaking the ball... and I think you can see where this is going...



Bulla-envelope with 11 plain and complex tokens inside. Near East, ca. 3700-3200 BC

Sumerian Cuneiform ~3300 BCE

It was a small step from imprinting the tokens on the outside of a ball to simply imprinting the tokens on a flat sheet of clay.



Administrative tablet with cylinder seal impression of a male figure, hunting dogs, and boars, Jamdat Nasr, Uruk III style.; 3100-2900 B.C. Now that there were no tokens stored with the contract, they needed symbols to represent numbers - the quantities in the contract.

And once you have numbers and nouns, it is a very small step to imagine adding other types of words like verbs, and representing entire sentences!

Photograph © www.metmuseum.org

Sumerian Cuneiform ~3300 BCE



The Sumerian script began with pictograms, but slowly evolved into the cuneiform that we recognize.

We don't know why they rotated the images 90 degrees at one point.

Cuneiform has the distinctive look that it has because the stylus had a triangle shaped tip ("cuneus" is Latin for wedge).

Cuneiform began as a logographic system, but evolved toward a syllabary over time.

We have so much information about cuneiform because (i) clay lasts a long time, and (ii) cuneiform was used until ~1000 BCE when Aramaic became the dominant language in the region.

Egyptian Hieroglyphics ~3000 BCE

The ancient Egyptians spoke a language called **Egyptian**. Current residents of Egypt speak a variety of Arabic called Egyptian Arabic. The two languages are not the same, though they are related, like cousins:

glyph	translit.	phonetic	glyph	translit.	phonetic
A	3	[7]	8	ķ	[h]
9	i	[i]	⊜	h	[x]
99	\\ у	[y]	÷	<u>h</u>	[ç]
السم	c	[٢]	<u> </u>	•– <i>s</i>	[s]
A	w	[w]		Š	[š]
J	b	[b]	۵	ķ	[q]
	p	[p]	Ŋ	k	[k]
×	f	[f]	ه	g	[g]
A	m	[m]	۵	t	[t]
~~~~	n	[n]	1	<u>t</u>	[ť ^y ], [c]
0	r	[r]	ŝ	d	[d]
	h	[h]	کر	₫	[ď ^y ], [j]



Egyptian hieroglyphs are unique in their time. The system itself was invented by the Egyptians. But we don't know if they independently invented the idea of writing, or if they got the idea from the Sumerians, and created their own.

Egyptian hieroglyphs, though pictures of objects, represent consonants in the Egyptian language. It is similar to an abjad... but with some complications.

#### Egyptian Hieroglyphics ~3000 BCE

Some glyphs represent two or three consonants. These consonant sequences are roots in the templatic morphology of Egyptian, so the glyphs can look more like logograms than an abjad:



Another complication is that sometimes a 2 or 3 consonant glyph will be followed by a 1 consonant root that repeats the final consonant. This can serve to disambiguate the word in cases where the first symbol could have more than one interpretation.



#### The Phoenicians (Canaan) ~1500 BCE

The Phoenicians (this is what the Greeks called them) get a lot of credit for spreading writing throughout the world. This is because they were major seatraders throughout the Mediterranean.



Egyptian hieroglyphs are most likely the source of the Phoenician script. You can see this through a series of links from Egyptian through earlier stages of writing in the area:

Egyptian hieroglyphic	Sinai script	Early Semitic	Name of letter
B	Å	¥	?aleph 'ox'
		4	bet 'house'
Y	Y	Ч	waw 'hook'
G	×	ĸ	kaph 'open hand'
,	~	щ	mem 'water'
J.	3	4	nahas 'snake'
$\langle \Phi \rangle$	0	0	Sajin 'eye'

#### The Phoenician script turns into Arabic and Hebrew and related abjads

#### History of the Arabic Alphabet

#### What are the historical origins of the Arabic writing system? This graphic helps examine the link between earlier alphabets and how they developed into the Arabic alphabet we know today, while also exploring how other languages from completely different language families use the Arabic alphabet as their primary writing system. Featured below is a selection of letters from each alphabet and how they have evolved throughout history.



#### Phoenician

The earliest recorded inscriptions of the Phoenician alphabet can be traced to c. 1200 BCE throughout modern-day Lebanon and Syria. The alphabet simplified over time, shifting from Egyptian hieroglyphics to the more stylized form seen above.

900 BCE

1200 BCE

•••••

200 CE

ڡ

0

2

**Syriac** 

The Syriac alphabet was a direct descendent of

the Aramaic alphabet, with its earliest inscriptions

dating to c. 200 CE. Syriac was primarily used as

the language of the church after the advent of

modern-day Arabic alphabet remains a highly

Christianity. The influence of Syriac on the

contested topic among scholars today.

#### Aramaic

The Arameans adopted the Phoenician alphabet beginning in c. 300 Bdz, and slowly adapted it into their own distinct alphabet. Aramics pread throughout the modern-day Middle East through the conquests of the Asyraian empire, who used the Arameans as scribes. After that, Aramaic was established as the lingua-franca for trade and commerce throughout the region.

0070 0097

#### Arabic

Between 400 CE and 700 CE, the Arabic alphabet developed from the Nabataean alphabet. Inscriptions throughout Greater Syria and the northern Arabian Peninsula display its influence today; however, the Arabic alphabet we use today was not developed until c. 700 CE.

400 CE - 700 CE

#### Nabataean

100 BCE

. . . . . . . . . . . . . . . .

The Nabataean alphabet was based on Aramaic and first appeared in c. 100 BCE in the ancient city of Petra, located in modern-day Jordan. Nabataean was written in two styles, monumental and cursive, with its cursive form displaying a clear connection to the modern Arabic alphabet.



.....

Healey, J. F., & Smith, G. R. (2009). A brief introduction to the Arabic alphabet: Its origins and various forms. London: Saqi. Gazai, O. Anabic-Persian Language Contact. The Semitic Languages: An International Handbook (pp. 1015–1021). Berlin, Germany: de Gruyter Mouton. Khan, A., Koka, N., & Anwar, S. Sociolinguistic impact of loamvords of Arabic origin on the Urdu language.(Report). Language In India, 13(1), 1006–1029. Mirdehghan, M., Persian, Urdu, and Pahto: A comparative orthographic analysis. Writing Systems Research, 2(1), 9–23.



c n

## .....

1200 CE

#### 

#### Persian

Arabic was first introduced to southwest and central Asia after a series of conquests by Arab Muslim dynasties in c. 650 CE. Persian adopted a modified version of the Arabic alphabet starting c. 900 CE.



Jiau

Urdu is spoken primarily in Pakistan, although it is also spoken in countries throughout southwest and central Asia. The Arabi Invasion of the Indian subcontineent in c. 1200 CE resulted in the languages meeting, leading Urdu to adopt a modified version of the Arabic alphabet closely related to the Persian alphabet developed earliet

**1500 CE** 

900 CE



Pashto

Pashto, a language closely related to Persian, adopted a modified version of the Arabic alphabet starting in c. 1500 CE. Pashto is one of Afghanistan's official languages and is also commonly spoken in Pakistan.

خلق روابط هادفة مع العالم العربي Inspiring Meaningful Connections with the Arab World

f QFINTL O QFINTL O QFINTL o qatar-foundation-international qfi.org

The Phoenician script also turns into Greek and related alphabets



#### A tree showing the spread of systems



### The Indus valley abugidas

Around 300 BCE, two scripts appeared in the Indus valley. One died out. It is called the "Indus Valley Script" because it is still undeciphered to this day. The other is Brahmi, which is the progenitor of all of the abugidas in South and Southeast Asia today!



#### Chinese characters ~ 1500 BCE

The Chinese script began as carvings on animal bone, often called the "Oracle bone script". You can see in this chart how the oracle bone carvings evolved into the Chinese characters that are in use today.

り	Ē	4	Ŧ	¥	赵	Y	H
人	男	女	子	夫	妻	Ξ	
rén	nán	nů	zi	fu	qī	wáng	kðu
person	man	woman	child	husband	wife	king	mouth
R	Ð	$\Diamond$	Θ	D	2	$\square$	Ð
目	耳	心	日	月	山	雨	田
mù	ěr	xīn	rì	yuè	shān	yů	tián
eye	ear	heart	sun	moon	mountain	rain	field
ים	ð.	સ્ત્ર	63	*	i la	÷	-
±	水	火	貝	大	小	F	下
tů	shui	huǒ	bèi	dà	xiǎo	shàng	xià
earth	water	fire	cowrie shell	big	small	above	below
ý	4	¥	Å	Л	Le bel	ş	Å
力	中	先	光	肉	出	Л	南
ĥ	zhōng	xiān	guāng	ròu	chū	dão	nán
strength	middle	first	bright	meat	to go out	knife	south
ላላ	\$	ત્રે	智	¥	¥	T.	<b>:</b>
北	東	X	好	豕	牛	馬	龜
běi	dōng	yòu	hảo	shi	niú	må	guī
north	east	also	good	pig	cow	horse	turtle



### China exports writing to Japan (~400 CE)

Japan first started writing with Chinese characters directly. But over time, developed its own writing systems based originally on the Chinese characters.

Japan has three writing systems in use today:

hiragana (syllabary) katakana (syllabary) kanji (Chinese characters)



# China exports writing to Korea (~400 CE), and then Korea invents its own system

Korea also first started writing using Chinese characters, again around 400 CE.

But in 1444, under the reign of King Sejong, a new system was invented, now called **Hangeul.** 

One interesting facet of Hangeul is that the shapes of some of the symbols were chosen based on articulatory phonetics (e.g., the position of the tongue). This may be the first writing system to try that!

7	77	2	ビ	TT	ビ	17
기역	쌍 기역	니은	디귿	쌍 디귿	리을	미음
giyeok	ssang giyeok	niən	digət	ssang digət	riəl	miəm
g/k	kk	n	d/t	tt	I	m
k/g	kk	n	t/d	tt	l/r	m
[k/g]	[k*]	[n]	[t/d]	[t*]	[l/r]	[m]
Ы	<b>A</b> A	入	*	0	ス	双
비읍	쌍 비읍	시옷	쌍 시옷	이응	지읒	쌍 지읒
biəp	ssang biəp	shiot	ssang shiot	iəng	jiət	ssang jiət
b/p	рр	s	SS	ng	j	jj
p/b	рр	s	SS	-ng	ch/j	tch
[p/b]	[p*]	[s]	[s*]	[Ø/-ŋ]	[tʃ/ʤ]	[ţʃ*]
え	ヨ	E	п	っ		
치읓	키읔	티읕	피읖	히읗		
chiət	kiuek	tiət	piəp	hiət		
ch	k	t	р	h		
ch'	k'	ť	p'	h		
[ʧ ^ʰ ]	$[k^h]$	[t ^h ]	$[p^h]$	[h]		

#### Mesoamerican glyphs (in modern day Mexico)

The Olmec civilization may have had writing around 900 BCE. The Cascajal block has 62 symbols (28 distinct) that may be written language. But we don't know for sure.

What we know for sure is that by 500 BCE the Mayan civilization did have a fully formed writing system. And we have deciphered about 80% of the system

Both of these civilizations died out without passing on their writing systems. So there are no modern writing systems that descend from Mesoamerica.





## Maya Hieroglyphs ~500 BCE

Maya glyphs are a mix of syllabary and logographic glyphs. The glyphs are presented in blocks. Here is an example of how the syllabary glyphs work.

Like Egyptian hieroglyphs and Chinese characters, glyphs can be logograms or phonetic symbols, and additional glyphs can be added to clarify what is intended.

In this example, the gray glyph can mean 'stone' or the syllable 'ku'.



https://www.mayaarchaeologist.co.uk/school-resources/maya-world/mayawriting-system/

### Undeciphered writing systems

#### Deciphering a writing system requires knowing the underlying language

In order to decipher a writing system, you must know the underlying language that it represents. If that is unknown, there is no way to decipher the system. The system only exists insofar as it represents a language.

Many of the undeciphered scripts in the world are undeciphered because we don't know the underlying language. (The one exception may be Rongorongo.)

The script called Linear A is a good example of this. It is from Crete. It was used around ~1600 BCE. But we do not know what language it represents.

Linear A probably gave rise to a script called Linear B, which we have deciphered - it represents Mycenaean Greek, and was used around ~1300 BCE, also on Crete. But Linear A remains a mystery.



### The Indus Valley script

The Indus Valley Civilization existed from 3300 BCE to 1300 BCE in Northern India and Pakistan. It was an incredibly sophisticated civilization, with dense cities with complex drainage and water systems. It dispersed over time, likely due to climate change.

There are over 4000 objects inscribed with the Indus Valley Script. There are over 417 distinct symbols on those objects.

Despite all of these examples, the Indus Valley Script remains undeciphered because we do not know the language spoken by the IVC.



#### Easter Island: Rapa Nui and Rongorongo

The indigenous language of Easter Island is Rapa Nui. There was a writing system in use on Easter Island as late as the 1860s that is (presumably) a representation of an older form of Rapa Nui. The writing is called Rongorongo.

The origin of rongorongo is unknown. Some carvings date to the 1600s CE, but because the only medium was wood, it is not clear if earlier carvings existed and have been lost.

The challenge with rongorongo is that we don't have much knowledge of older forms of Rapa Nui, and we don't have any key to link some set of the symbols to known words (like the Rosetta stone for Egyptian hieroglyphs). Some scholars even debate whether it is a fully fledged writing system.

