

Alphabets II & Abugidas I

Lecture 7

February 28, 2022

Quiz 2

■ Options:

- During class: At the start of class (10:15am)
- Outside of class: Keep previous format of a 6 hour window on the day of the quiz

■ Survey Results

- Outside of class: Keep previous format of a 6 hour window on the day of the quiz

Attempts: 53 out of 53

When would you like the following two quizzes (Quiz 2 and Quiz 3) in this course to take place?

During class: At the start of class (10:15am)	5 respondents	9 %	 ✓
Outside of class: Keep the Quiz 1 format of a 6 hour window on the day of the quiz	48 respondents	91 %	

Quiz 2

- Similar format to last time
- Everything we've covered since Quiz 1
 - Abjads
 - Aegean / Mediterranean scripts
 - Alphabets
 - Abugidas (this lecture only)
 - + associated sections in the book

Quiz 2

- You will have to choose one of Hebrew or Arabic to study for the quiz
 - Script tables & transliteration practice words are now posted
 - Answers will be posted March 3
 - You will not have to learn the system by heart
 - However, doing your best to learn one means you will be able to answer these questions faster
 - I also expect you to understand / apply your knowledge of matres lectionis
 - Questions will be abjad → romanization (you won't have to write Hebrew / Arabic!)
 - Words will be (sufficiently) vowelled

Textbook chapters

■ Chapter 7: Semitic

- skip the discussion of Middle Persian in §7.3.3 (par. 1 on p. 124)—we'll revisit this later
- skip the short note on Hatran script and Sogdian script (p. 124–5)
- I won't ask specific questions about historical sound changes in Hebrew (§7.4.2), but you may still find reading this section helpful
- skip the bit about vowel points and the name of God in §7.4.5
- skip the Arabic numerals in §7.5.7

■ Chapter 8: The Greek Alphabet

- skip Armenian and Georgian in §8.6.3–4

Textbook chapters cont...

- Chapter 9: The Roman Alphabet
 - Skip §9.7 for your quiz studying, but we'll be talking about this after the quiz in lecture
- Chapter 11: The Indian Abugida
 - skip the brief outlines of Brahmic daughter scripts in §11.3.4
 - don't worry about the romanization details in §11.4.1
 - You can stop reading this chapter on p. 218! We are skipping the finer details of Sanskrit, and we haven't gotten to other scripts yet. This means:
 - Skip §11.4.2.3–7 (p. 218–20)
 - Skip 11.5–8 for your quiz studying, but we'll also cover most of this material in lecture this week

Today...

Alphabets, pt. II

- Roman alphabet
 - Etruscan
 - Origins
 - Loss & gain of letters
- Adaptation of alphabets

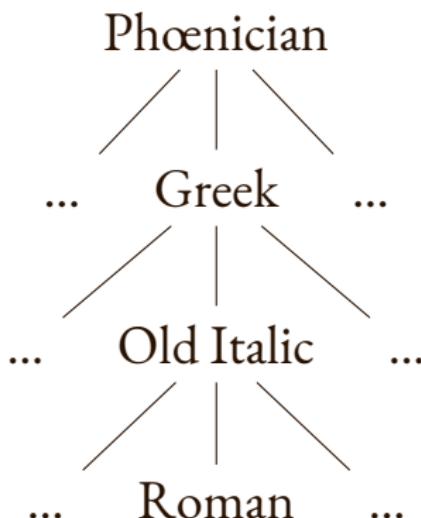
Abugidas

- Ethiopic (Ge'ez)
- Indic Writing Systems
 - Indus Script
 - Brāhmī & Kharoṣṭhī
 - Connection to Buddhism
- Devanāgarī
- Writing clusters

Roman alphabet

Roman alphabet

- Also called the Latin Alphabet
 - Descended from the Greek alphabet, via Etruscan
 - Underwent further modification, and has been widely adapted for many languages
 - Earliest Old Italic inscriptions: 7th century BCE
 - Used alongside Greek in many inscriptions



Etruscan language

■ Language isolate

- Not fully deciphered, despite some recent claims

- Borrowed the Greek alphabet, but wrote it right-to-left

- Etruscans were conquered by Romans

- Romans adapted their script to write Latin
 - Always wrote it left-to-right

Euboean Greek	Model Etruscan	Archaic Etruscan	Late Etruscan	Latin	Phonetic Value
Α Α	Α	Α	Α	A	[a]
Β	Φ			B	[b]
Ϲ	Γ))	C G	[k]
Δ	Δ			D	[d]
Ϛ Ε	Ϛ Σ	Ϛ	Ϛ	E	[e]
Ϛ Φ	Ϛ Φ	Ϛ	Ϛ	F	[w]
Ϛ Ι	Ϛ Ι	Ϛ	Ϛ	(Z)	[z]
Ϛ Η	Ϛ Η	Ϛ	Ϛ	H	[h]
Ϙ Ϙ ϖ	Ϙ Ϙ ϖ	Ϙ Ϙ ϖ	Ϙ Ϙ ϖ		[tʰ]
Ϛ	Ϛ	Ϛ	Ϛ	I	[i]
Κ	Ϙ	Ϙ	Ϙ	K	[k]
Ӆ	Ӆ	Ӆ	Ӆ	L	[l]
Ϻ Μ	Ϻ Μ	Ϻ	Ϻ	M	[m]
ϻ Ν	ϻ Ν	ϻ	ϻ	N	[n]
Ϛ Σ	Ϛ Σ	Ϛ	Ϛ	O	[s]
Ϛ Ο	Ϛ Ο	Ϛ	Ϛ	P	[o]
Ϛ Ρ	Ϛ Ρ	Ϛ	Ϛ		[p]
Ϛ Ω	Ϛ Ω	Ϛ	Ϛ	Ϛ	[s̪]
Ϛ Ρ	Ϛ Ρ	Ϛ	Ϛ	Q	[q]
Ϛ Σ	Ϛ Σ	Ϛ	Ϛ	R	[r]
Ϛ Τ	Ϛ Τ	Ϛ	Ϛ	S	[s]
Ϛ Υ	Ϛ Υ	Ϛ	Ϛ	T	[t]
Ϛ Χ	Ϛ Χ	Ϛ	Ϛ	V	[u]
Ϛ Φ	Ϛ Φ	Ϛ	Ϛ		[ks]
Ϛ Ψ	Ϛ Ψ	Ϛ	Ϛ	Ϛ	[pʰ]
(88)			8		[kʰ]
(88)					[f]

Etruscan Abecedary

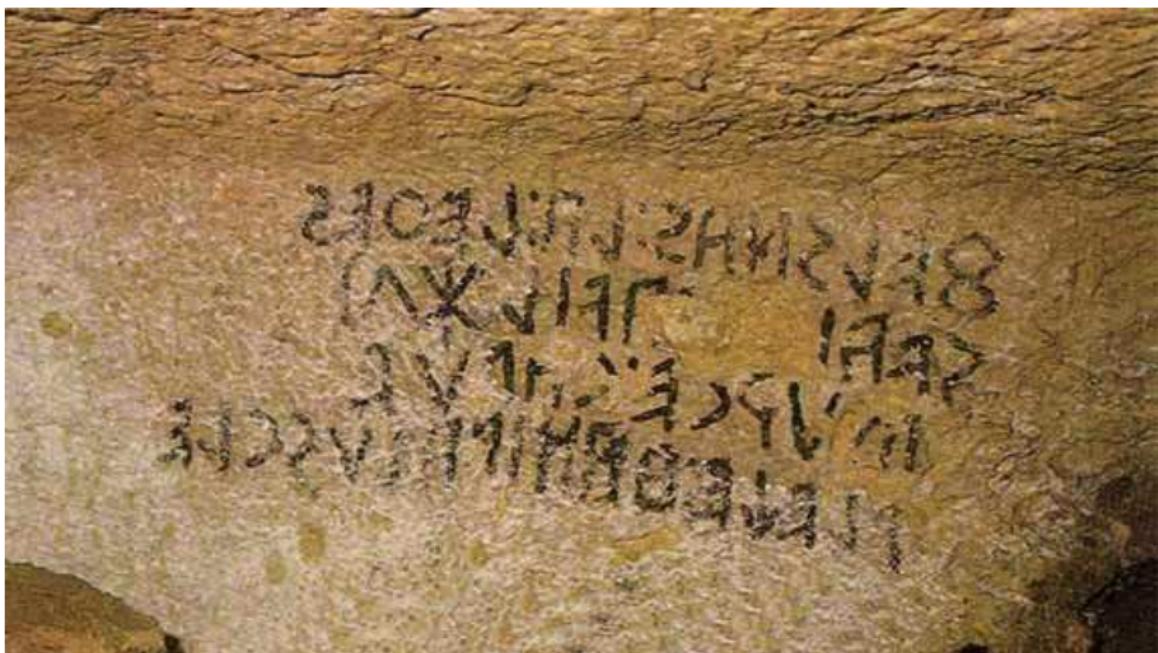
- The Marsiliana Abecedary is one of the earliest such sign lists that we find in an Old Italic alphabet
 - **abecedary**: alphabet list
 - Dated to around 700 BCE
 - At this point the Etruscan alphabet still had some Greek graphemes that the Roman alphabet does not (digamma, san, qoppa)



ABIDE IN MY SIGHT AND FOLLOW ME

Roman alphabet

○○●○○○○○○○○○○



Etruscan inscription TLE 890

felsnas : la : leθes

svalce : avil CVI

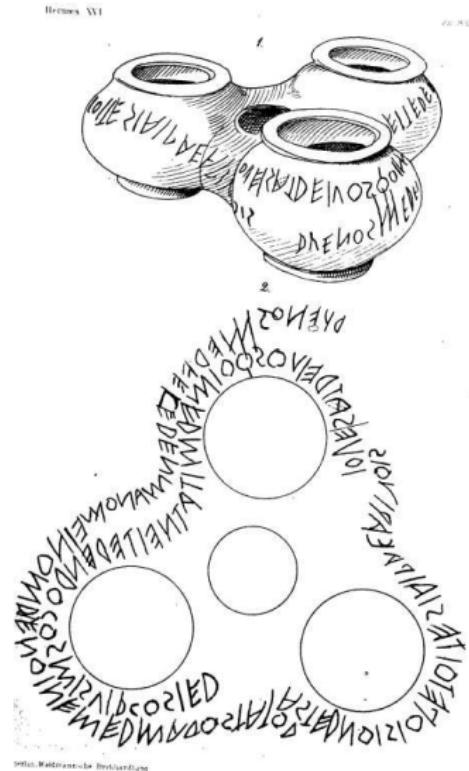
murce : capue

tlexe : hanipaluscle

Early Roman alphabet

- Roman alphabet borrowed from Etruscan
 - Duenos inscription: one of the earliest inscriptions in Old Latin

IOVESATDEIVOSQOIMEDMITAT
NEITEDENDOCOSMISVIRCOSIED
ASTEDNOISIOPETOITESIAIPAK
ARIVOIS
DVENOSMEDFECEDEN
MANOMEINOMDVENOINEMED
MALOSTATOD



Early Roman alphabet

■ Other graphemes

- /ks/ → ⟨X⟩ based on West Greek
 - ⟨CS⟩ or ⟨KS⟩ not used
 - ⟨Z⟩ originally discarded
 - No /z/ phoneme in Latin
 - Later reintroduced to write Greek loans

Names of the letters

- Greek letters taken directly from their Phoenician sources
 - $\langle A \rangle$ *aleph* → *alpha*
 - $\langle B \rangle$ *bet* → *beta*
 - $\langle D \rangle$ *dalet* → *delta*
 - ...
- Romans gave new names to the borrowed graphemes
- Vowels were named for their sounds:
 - $\langle A \rangle$ /a/ → *a*
 - $\langle E \rangle$ /e/ → *e*
 - $\langle I \rangle$ /i/ → *i*
 - $\langle O \rangle$ /o/ → *o*
 - $\langle U \rangle$ /u/ → *u*
- Consonants added /e/ before or after:
 - $\langle B \rangle$ /b/ → *be*
 - $\langle P \rangle$ /p/ → *pe*
 - $\langle M \rangle$ /m/ → *em*
 - $\langle X \rangle$ /ks/ → *ex*

Modern names

- English uses the Roman names, but sound changes have altered the vowels

- /a:/ is /ej/
- /e:/ is /i/
- /i:/ is /aj/
- /o:/ is /o/
- /u:/ is /ju/

- Likewise for consonants

- /be:/ is /bi/
- /te:/ is /ti/
- ...

⟨Z⟩

- /zəd/ or /zi/?
- ⟨Z⟩ given Greek name /zeta/ when borrowed back into Roman alphabet
- This carried into English as /zəd/
 - Used in most of English-speaking world
- American dictionary author Noah Webster suggested /zi/, for uniformity with other letters
 - /ze:/ would have been the Roman name for ⟨Z⟩, if they had kept it
 - Latin /ze:/ would have become English /zi/

Lost letters

- The English alphabet has had a couple of other letters at earlier points in its history
- These appear in Old & Middle English texts

- | | |
|---|---|
| <ul style="list-style-type: none">■ Thorn: þ, þ■ Borrowed from rune <i>thurs</i> (þ)■ Represented /θ/ | <ul style="list-style-type: none">■ Wynn: þ, þ■ Borrowed from the rune of the same name (þ)■ Represented /w/, previously written ⟨uu⟩■ Later removed, ⟨uu⟩ → ⟨w⟩ |
|---|---|

“Ye”

- In Early Modern English, ⟨the⟩ was written ⟨þe⟩ (with a thorn)
- The ⟨þ⟩ and ⟨y⟩ look very similar in Blackletter script
- Once thorn was dropped from the alphabet, ⟨þe⟩ in older writings was misread as ⟨ye⟩
 - These instances of “the” are therefore often pronounced like the pronoun “ye” for an old-timey feel
 - This was never its pronunciation, and “ye” was never actually used as “the”



Common manuscript rendition of ⟨þe⟩,
with superscript ⟨e⟩

Adaptation

- Roman is the basis of most of the IPA
- Latin-based orthographies are used in virtually all language families
- Alphabets are extended mostly using diacritics, polygraphs
 - Example: Postalveolars in Roman alphabet languages

Postalveolars

- Latin had no postalveolar consonants /ʃ, tʃ/
 - Many of the languages using the Roman alphabet do, and needed to devise ways to write these sounds
 - Some orthographies represent them with digraphs or polygraphs
 - (Graphemes in red only correspond to the sound in certain vowel environments)

	<i>English</i>	<i>French</i>	<i>German</i>	<i>Italian</i>	<i>Basque</i>
/s/	$\langle s \rangle, \langle (s)c \rangle$	$\langle s \rangle, \langle c \rangle$	$\langle s \rangle$	$\langle s \rangle$	$\langle s \rangle$
/ʃ/	$\langle sh \rangle, \langle si \rangle, \langle ti \rangle$	$\langle ch \rangle$	$\langle sch \rangle$	$\langle sc \rangle$	$\langle x \rangle$
/tʃ/	$\langle ch \rangle$	$\langle tch \rangle$	$\langle tsch \rangle$	$\langle c \rangle$	$\langle tx \rangle$

	<i>Hungarian</i>	<i>Polish</i>
/s/	⟨sz⟩	⟨s⟩
/ʃ/	⟨s⟩	⟨sz⟩

Postalveolars

- Others created new graphemes using diacritics
 - (Graphemes in red only correspond to the sound in certain vowel environments)

	<i>Czech</i>	<i>Turkish</i>	<i>Romanian</i>
/s/	⟨s⟩	⟨s⟩	⟨s⟩
/ʃ/	⟨š⟩	⟨ş⟩	⟨ş⟩
/tʃ/	⟨č⟩	⟨ç⟩	⟨c⟩

N'Ko

Manding languages



N'ko

- The Manding languages, spoken in West Africa
 - Bambara, Mandinka, Maninka, Dyula, ...
- In 1949, a Maninka writer, Solomana Kante, devised the N'ko script for these languages
 - N'ko = “I speak”
 - Felt that the lack of an indigenous writing system for his language made his people “cultureless”
 - The system was popular and spread through the Manding languages
- Inspired by Arabic, but an alphabet!
 - Grapheme inventory designed for these languages! (not adapted / transplanted from Arabic)
 - Graphemes are connected like Arabic, but shapes are unrelated
 - Still written L ← R

Roman alphabet

oooooooooooooooo

N'Ko

○○●○○○○

Ge'ez

○○○○○○○○○○

Indic Scripts

○○○○○○○○○○○○○○○○○○○○○○○○○○○○○○

Devanāgarī

○○○○○○○○○○○○

N'ko



Looks sort of familiar...

- Like Arabic: connected free graphemes + diacritics
- Not like Arabic: what the graphemes represent

የፋይል ቤትና የፋይል ቤትና ቤትና
ይህንን ቤትና ቤትና

/ fdàfíná làdòkòyàli gbrèmàtè:li /
/ kzabulāka bẽmàkā nù /

Free & bound graphemes

- Both vowels and consonants are represented with free graphemes...

F I
 /b/ /a/

ঁ ঊ
 /gb/ /e/

ବ ଓ ତ ଯ
 /t/ /e/ /r/ /i/

ଫ “river”
 /ba/

ଅଙ୍ଗ “oak”
 /gbe/

ଯତୋବ “friend”
 /teri/

Free & bound graphemes

- ...so then what are these? (*Video: N'ko diacritics*)

$\bar{\text{o}}$	$\tilde{\text{o}}$	$\dot{\text{o}}$	o	short
/é/	/è/	/ě/	/ê/	
$\grave{\text{o}}$	$\hat{\text{o}}$	$\acute{\text{o}}$	$\hat{\text{o}}$	long
/é:/	/è:/	/ě:/	/ê:/	
high tone	low tone	rising tone	falling tone	

- ...and this?

$\ddot{\text{o}}$
/ě/
nasalization

Roman alphabet

ooooooooooooooo

N'Ko

oooooo●

Ge'ez

oooooooooooo

Indic Scripts

oooooooooooooooooooooooooooo

Devanāgarī

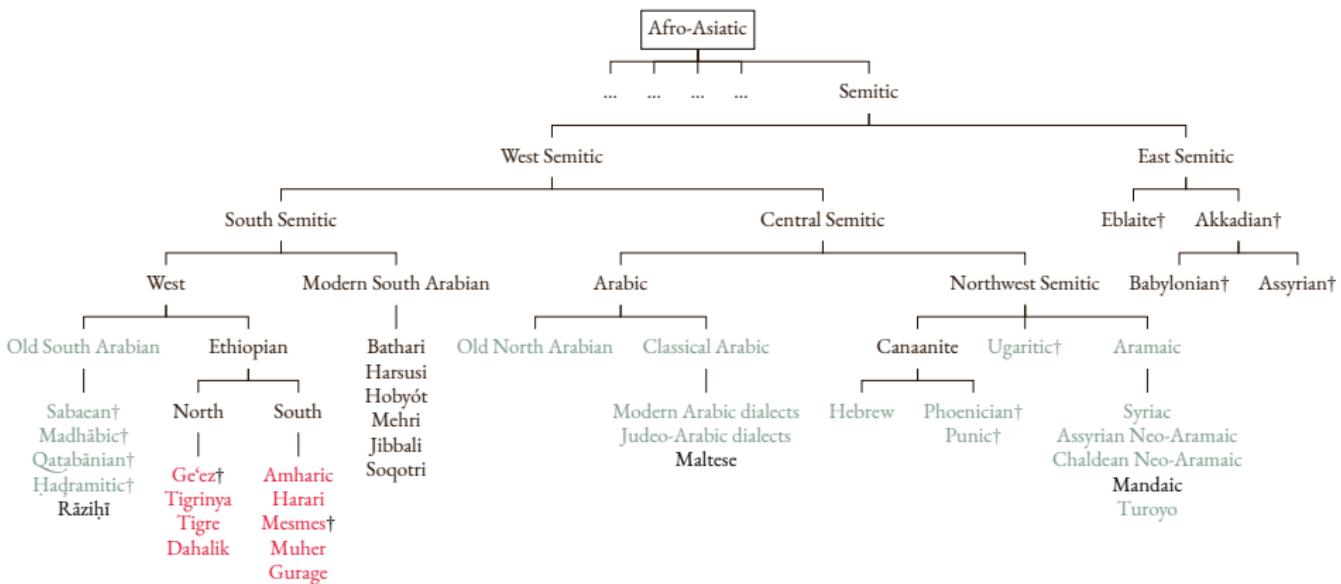
oooooooooooooooo

N'ko

/ɔ/	/o/	/u/	/ɛ/	/i/	/e/	/a/
Ω	¤	¤	ʌ	Y	o	I
/r/	/d/	/tç/	/dz/	/t/	/p/	/b/
t	¤¤	l	Ɣ	b	ঁ	F
/m/	/l/	/k/	/f/	/gb/	/s/	/rr/
Δ	¶	¶	ڻ	ڻ	□	ٿ
/n/		/j/	/w/	/h/	/n/	/ŋ/
ঠ		ɸ	ঘ	঱	ঠ	ঁ

Ge'ez

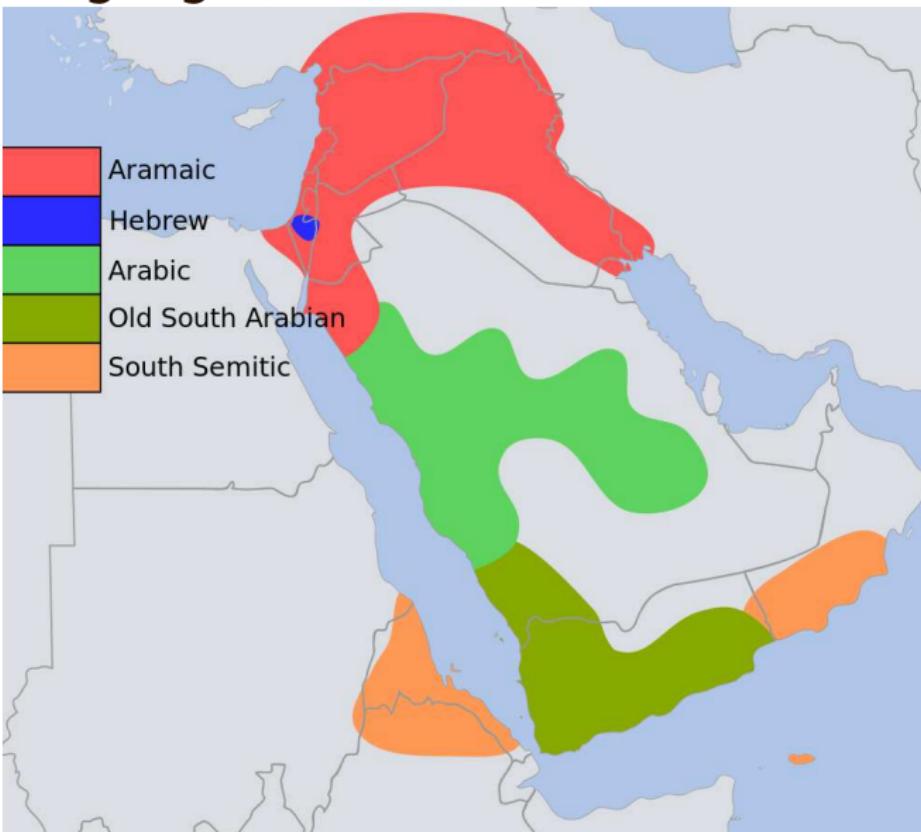
Semitic languages



*Semitic languages written in an abjad.
Semitic languages written in an abugida.*

† – Dead language.

Semitic languages, ~100 AD



A sister of Phoenician

- We mentioned the **Proto-Sinaitic (Proto-Canaanite)** script a few weeks back
 - Developed through acrophony from Egyptian hieroglyphics
 - ~1500–1200 BC
 - Precursor of Phoenician abjad
- It was also the precursor of another script
 - North Semitic: Phoenician (from ~1200 BC)
 - South Semitic: **South Arabian** diverges ~1400 BC



South Arabian

- South Arabian script, also called Old South Arabian or Epigraphic South Arabian
 - **Abjad**, used 900 BC – 700 AD
 - Usually written R-L, but can also be written L-R (with graphemes flipped)
 - ~10,000 inscriptions, most in stone,
~1000 on wooden cylinders
 - Used to write the South Arabian languages in the far south of the Arabian peninsula
- (There is also a related Old North Arabian script)



Roman alphabet

ooooooooooooooo

N'Ko

ooooooo

Ge'ez

oooo●oooo

Indic Scripts

oooooooooooooooooooooooooooo

Devanāgarī

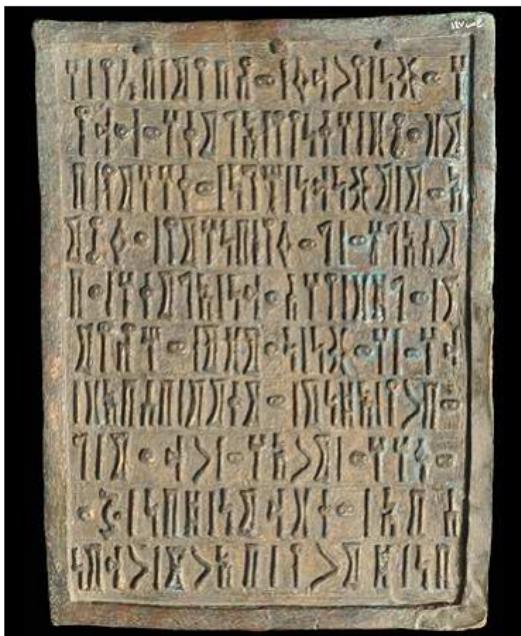
oooooooooooooooooooo

Epigraphic South Arabian

		Phœ.		ESA		Phœ.		ESA
b		ബ	ബ	ܒ		ബ	ബ	ܒ
l		ി	ി	ܵ		ി	ി	ܵ
b̄		ി	ി	ܴ		ി	ി	ܴ
m		ഓ	ഓ	ܸ		ഓ	ഓ	ܸ
q		ഔ	ഔ	ܹ		ഔ	ഔ	ܹ
w		ി	ി	ܺ		ി	ി	ܺ
ſ		ഉ	ഉ	ܻ		ഉ	ഉ	ܻ
r		ാ	ാ	ܾ		ാ	ാ	ܾ
b̄̄		ാ	ാ	ܴܶ		ാ	ാ	ܴܶ
t̄		ം	ം	ܳ		ം	ം	ܳ
s̄		ം	ം	ܲ		ം	ം	ܲ
k̄		ം	ം	ܰ		ം	ം	ܰ
n̄		ം	ം	ܱ		ം	ം	ܱ
b̄̄̄		ം	ം	ܴܶܶ		ം	ം	ܴܶܶ

Epigraphic South Arabian

- Earliest Epigraphic South Arabian texts are actually found in modern-day Eritrea and Ethiopia
 - Kingdom of D'mt (ዳምጥ, କୋମ୍ପତ୍), 900–400 BC
 - Along with other nearby kingdoms, became the Kingdom of Aksum around 100 AD
- Around the time of the Aksumite Kingdom, Ge'ez **abjad** appears
 - Written **left to right**
 - Shapes of graphemes virtually identical to the modern Ge'ez script



Roman alphabet

ooooooooooooooo

N'Ko

ooooooo

Ge'ez

oooooooo●oooo

Indic Scripts

oooooooooooooooooooooooooooo

Devanāgarī

oooooooooooooooooooo

Ge'ez abjad

	Ge'ez	Phœ.	ESA	Ge'ez	Phœ.	ESA	Ge'ez
b	ብ	ብ	ብ	ብ	w	ወ	ወ
l	ል	ል	ል	ል	z	ዕ	ዕ
ḥ	ሮ	ሮ	ሮ	ሮ	y	የ	የ
m	ም	ም	ም	ም	d	ዳ	ዳ
ś	ሮ	ሮ	ሮ	ሮ	g	ገ	ገ
r	ሩ	ሩ	ሩ	ሩ	t	ተ	ተ
s	ሮ	ሮ	ሮ	ሮ	ṗ	ፖ	ፖ
q	ቆ	ቆ	ቆ	ቆ	s̥	ሮ	ሮ
b	በ	በ	በ	በ	d̥	ዶ	ዶ
t	ተ	ተ	ተ	ተ	f	ፈ	ፈ
ḥ	ሮ	ሮ	ሮ	ሮ	p	ጠ	ጠ
n	ነ	ነ	ነ	ነ	s̥ʒ	ሮ	ሮ
?	፳	፳	፳	፳	t̥	ጥ	ጥ
k	ከ	ከ	ከ	ከ	z̥	ዘ	ዘ

Ge'ez abugida

- Original shapes of the letters were later marked with diacritics to indicate vowels
- These diacritics fused with the consonants, resulting in the modern **abugida** system
- Ge'ez script is used to write Amharic, Tigrinya, and the other Ethiopian Semitic languages spoken in Ethiopia and Eritrea



Abugida

- System where each grapheme indicates a V or CV syllable
 - Base grapheme has an **inherent vowel**
 - Other vowels are indicated through diacritics or modification of base grapheme shape
- Abugida is Ge'ez name for the script: አቡጋዳ (’äbugida)
- In Ge'ez script, <Ci> graphemes are used for consonants without vowels: /Ci/ or /C/

	/ə/	/u/	/i/	/a/	/e/	/ɪ/	/o/
/s/	ሰ	ሸ	ዴ	ሳ	ሸ	ዴ	ሬ
/b/	በ	ቡ	ቡ	ባ	ቡ	ብ	ቦ
/k/	ከ	ኻ	ኻ	ካ	ኻ	ኻ	ኻ
/z/	ዘ	ዘ	ዘ	ዛ	ዘ	ዘ	ዘ

Systematic, but complex

- Multiple related strategies for indicating a vowel

	/ə/	/u/	/i/	/a/	/e/	/ɪ/	/o/
/w/	ወ	ወ	ዊ	ዊ	ዊ	ወ	ዊ
/ɿ/	ዕ	ዕ	ዊ	ዊ	ዊ	ዕ	ዊ
<hr/>							
/m/	ወ	ወ	ዊ	ዊ	ዊ	ጥ	ጥ
/ś/	ወ	ወ	ዊ	ዊ	ዊ	ሮ	ሮ
<hr/>							
/p/	ጥ	ጥ	ጥ	ጥ	ጥ	ጥ	ጥ
/g/	ገ	ገ	ገ	ገ	ገ	ገ	ገ
/k'/	ቁ	ቁ	ቁ	ቁ	ቁ	ቁ	ቁ
<hr/>							
/r/	ሩ	ሩ	ሩ	ሩ	ሩ	ሩ	ሩ
/f/	ፈ	ፈ	ፈ	ፈ	ፈ	ፈ	ፈ

Addition of graphemes

- Graphemes were added to the Ge'ez script to write loanwords and words in related languages
- A line above (originally a diacritic?) is usually used to mark either **palatalization** or **spirantization**

/t/	ተ	/tʃ/	ቃ	/k/	ከ	/χ/	ቂ
/t'/	ጠ	/tʃ'/	ቃ	/k'/	ቁ	/χ'/	ቁ
/d/	ደ	/dʒ/	ጋ	/b/	በ	/v/	ቊ
/s/	ሮ	/ʃ/	ሻ				
/z/	ዘ	/ʒ/	ሻ				
/n/	ኅ	/ɳ/	ኅ				

Indic Scripts

Devanāgarī **शिवो रक्षतु गीर्वाणभषारसास्वादतत्परान्**

Gurmukhi ਸਿਵੇ ਰਕਸਤੁ ਗੀਰਵਾਣਭਸਾਸ੍ਵਾਦਤਤਪਰਾਨ्

Bengālī শিবো রক্ষতু গীর্বাণভাষারসাস্বাদতত্পরান্

Gujarātī શિવો રક્ષતુ ગીર્વાણભાષારસાસ્વાદતત્પરાન્

Tibetan གී ར් ཛ ད བ ཉ དྷ ང ཉ ང ཉ ང ཉ

Odiā ଶିବୋ । ରକ୍ଷତୁ ଗୀର୍ବାଣଭାଷାରସାସ୍ଵାଦତତ୍ପରାନ୍

Telugu శివో రక్షతు గీర్వాణభాషారసాస్వాదతత్పరాన్

Kannada ಶಿವೋ ರಕ್ಷತು ಗೀರ್ವಾಣಭಾಷಾರಸಾಸ್ವಾದತತ್ಪರಾನ್

Sinhala ශිවෝ රක්ෂතු ගීර්ඩ්‌බාහ්‌රසා‌ස්වාදත්පරාන්

Malayālam ശിവോ രക്ഷതു ഗീര്വാണഭാഷാരസാസ്വാദതത്പരാന്

Tamil ஶிவோ ரக்ஷது கீர்வாணபாஷாரஸாஸ்வாததத்பரந்

Burmese ရိပိဝါ ရက္ခတု ဂိုဏ်သာမာရသာသွေအတတ္ထရန်

Thai ສີໂຮກເຖິງຄົວລາວພາເມືອນກາສາສຸວັດທະນາບຸ

Lao ສີໂຮກສະຕິວານະພາສາກະສາວາທະຕິປະໄານ

Khmer ພີ່ນົກງູດຸກົງດາວັດຫຍາສູງສູງ

Balinese ଶିବୋ ରକ୍ଷତୁ ଗୀର୍ବାଣଭାଷାରସାସ୍ଵାଦତତ୍ପରାନ୍

Javanese චිඩ්‌යානා මායා ගායා ගායා මායා ගායා ගායා ගායා ගායා ගායා

Sundanese චිඩ්‌යානා මායා ගායා ගායා ගායා ගායා ගායා ගායා

Buginese චිඩ්‌යානා මායා ගායා ගායා ගායා ගායා ගායා

(Sanskrit trans.) <śivo rakṣatu gīrvāṇabhāṣārasāsvādatatparān>

“May Shiva bless those who take delight in the language of the gods.”

Language families

■ Indo-Aryan

- Major branch of Indo-European family
- Dozens of languages
 - Hindi/Urdu, Bengali, Punjabi, Marathi, Gujarati, Bhojpuri, Oriya, Sindhi, Sinhala, Nepali, Assamese...
- Over a billion speakers

■ Dravidian

- Separate unrelated language family (not Indo-European!)
- Dozens of languages
 - Telugu, Tamil, Kannada, Malayalam, Brahui, Kurukh, Tulu...
- A few hundred million speakers

Roman alphabet

oooooooooooooo

N'Ko

oooooooo

Ge'ez

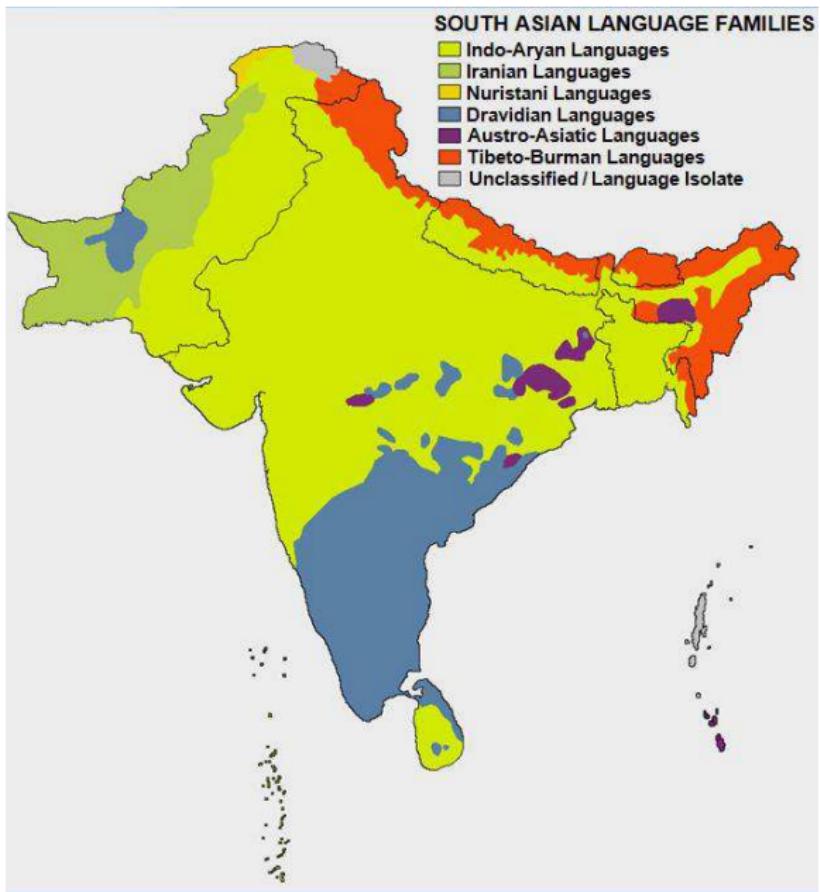
oooooooooooo

Indic Scripts

ooo●oooooooooooooooooooooooooooo

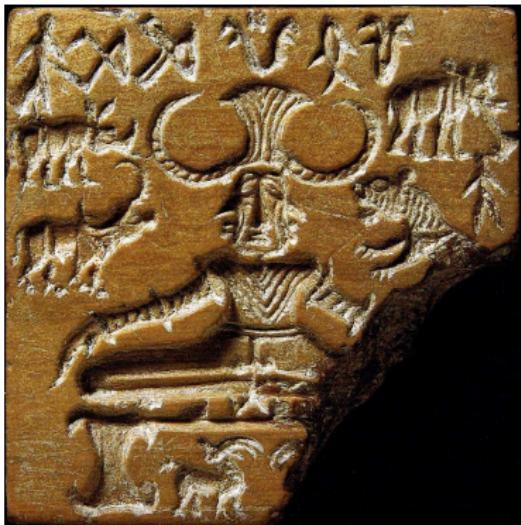
Devanāgarī

oooooooooooo



Indus Valley civilization

- 3600–1900 BCE
- Modern-day Pakistan
- Indus script
 - Inscriptions found from ~2500 BCE
 - Mostly pressed clay seals
 - Possibly lost texts on organic material
 - *Unknown language*
 - *Undeciphered script*



Seal from Mohenjodaro, depicting *paśupati*
“Lord of Animals”

Mohenjodaro



- One of the largest Indus Valley settlements
- Modern-day southern Pakistan

Roman alphabet

oooooooooooooo

N'Ko

ooooooo

Ge'ez

oooooooooooo

Indic Scripts

oooooooo●oooooooooooooooooooo

Devanāgarī

oooooooooooooooo

Indus Valley seals



Indus script

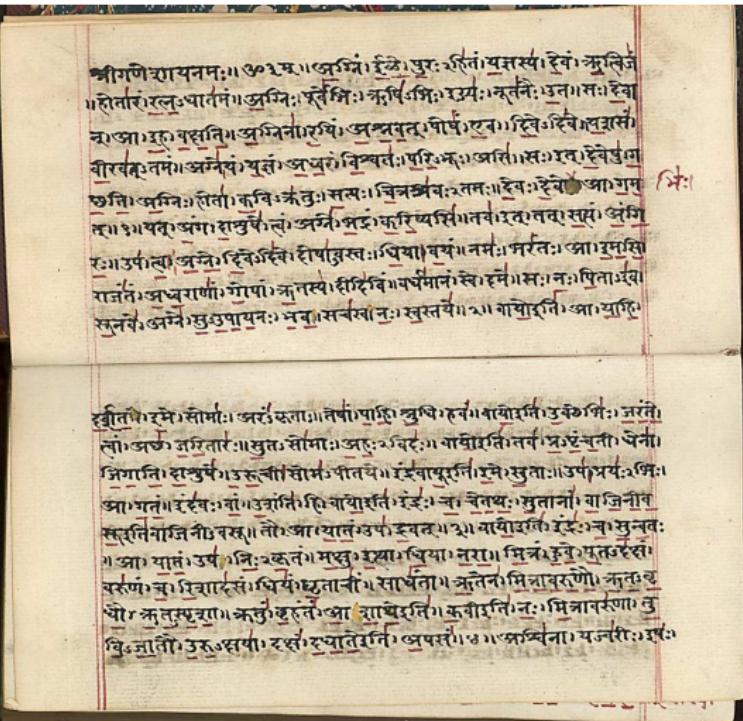
- Pictographic appearance
- Probably moraic with some morphograms
 - Too many signs to be purely phonemic or moraic
 - Too few to be morphographic
 - Average inscription length suggests this too
- Language likely Dravidian
 - Found in pockets in the north
 - Hypothesized southern origin, then forced north
- Stimulus diffusion

Brāhmī & Kharoṣṭhī

- ~3rd century BCE, we find more widespread writing systems
- Two scripts: Brāhmī and Kharoṣṭhī
- Most preserved texts written in stone
 - Written in Prakrit (descendent of Sanskrit)
 - Preachings of Buddhist moral values by Emperor Aśoka

Sanskrit

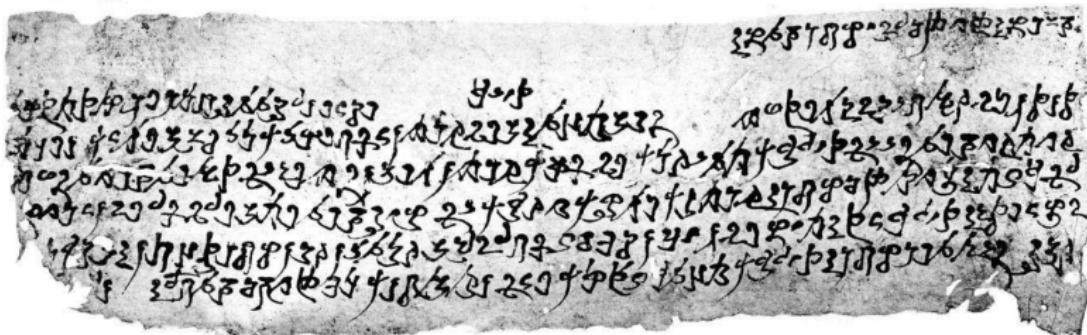
- Sacred language of Hinduism, philosophical language of Buddhism
- Spoken ~2000–600 BC
- Continues to be learned and used as a written liturgical language for Hinduism, Buddhism
- Some efforts to revive it as a spoken language



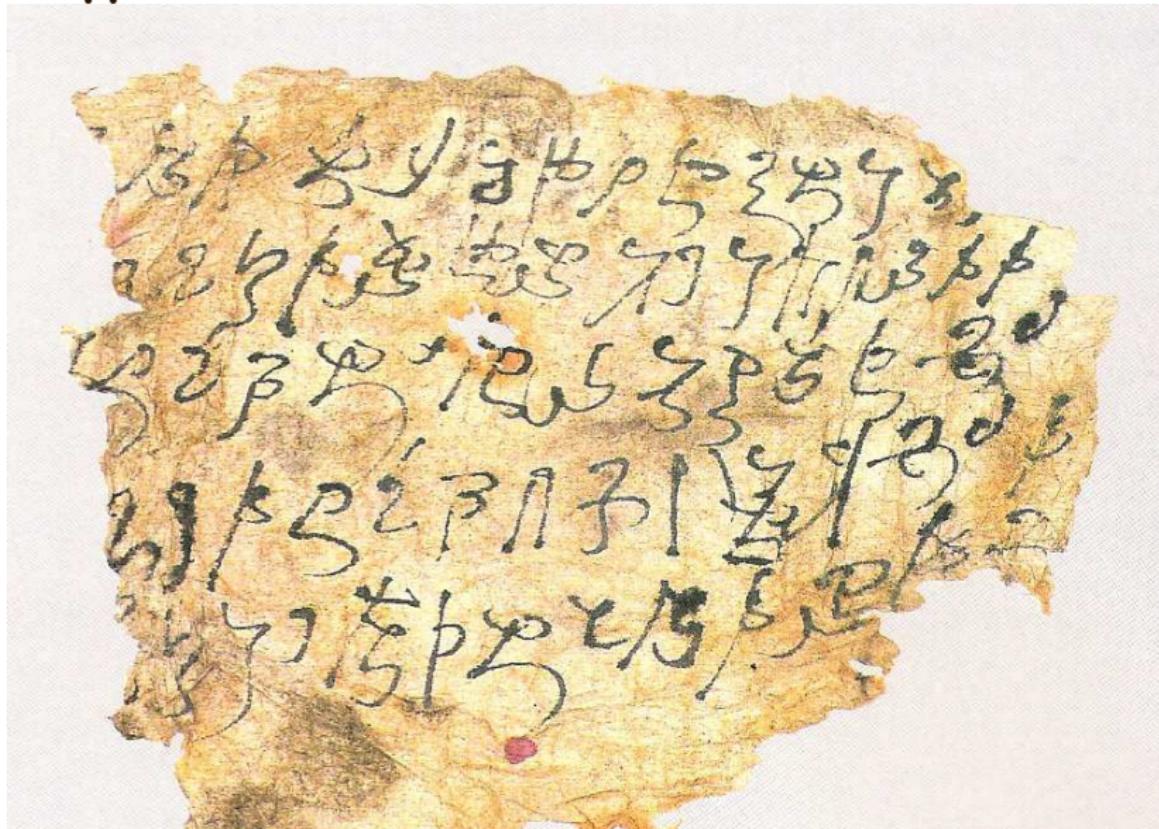
Vedic (pre-Classical) Sanskrit, written ~early 1800s

Kharoṣṭhī

- Somewhat cursive in appearance
- Written **right-to-left**
- No vowel length distinctions written
- Replaced by Brāhmī

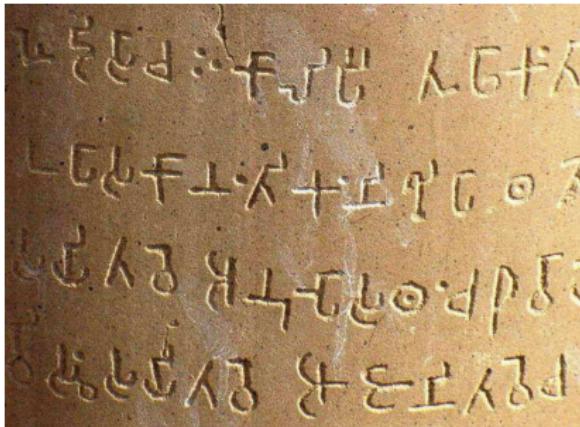


Kharoṣṭhī



Brāhmī

- Not at all cursive
- Written **left-to-right**
- Adapted to many other languages and diverse scripts
 - Collectively referred to as “Brahmic scripts”
 - Mostly retained the same number & structure of graphemes due to conservative influence of Sanskrit



Roman alphabet

ooooooooooooooo

N'Ko

ooooooo

Ge'ez

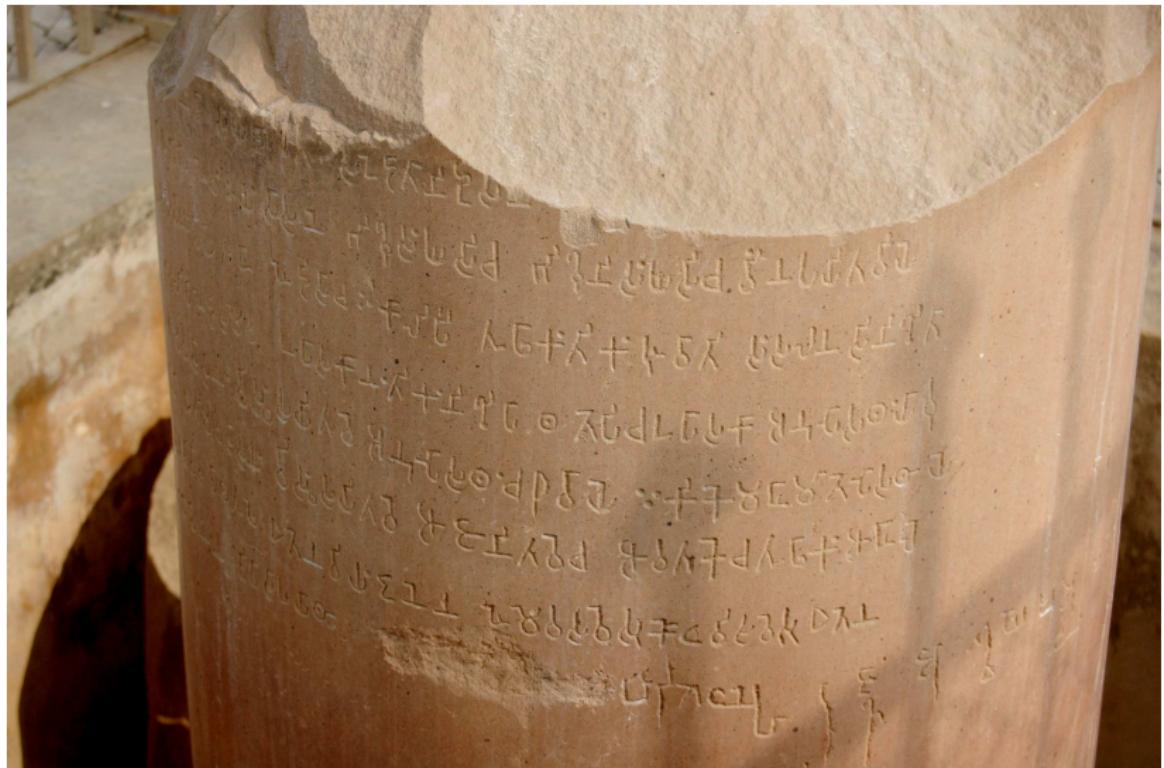
oooooooooooooo

Indic Scripts

oooooooooooo●oooooooooooo

Devanāgarī

oooooooooooooooo



Pillar of Sarnath, ~250 BC

Decipherment of Kharoṣṭhī

- Bilingual coins from Indo-Greek Kingdom (180 BC – 10 AD)
 - Hellenistic kingdom covering modern Afghanistan & Pakistan
- Coins had inscriptions in Greek and Pāli
 - Pāli: a Prakrit language (desc. of Sanskrit) written in Kharoṣṭhī
- James Prinsep (English scholar, 1799–1840)
 - Used these inscriptions to decipher Kharoṣṭhī
 - Also deciphered Brāhmī!



Indo-Greek pedestal: Bodhisattva with Corinthian columns and Kharoṣṭhī text, ~380 BC

Decipherment of Kharoṣṭhī



Silver bilingual tetradrachm of King Menander I (155–130 BC)

- Greek: ΒΑΣΙΛΕΩΣ ΣΩΤΗΡΟΣ ΜΕΝΑΝΔΡΟΥ
 - ⟨basileos soteros menandrou⟩
- Pāli: පෙර ලලක්ෂණ මහාරාජා
 - ⟨maharaja tratarasa menadrasa⟩
- “Of Saviour King Menander”

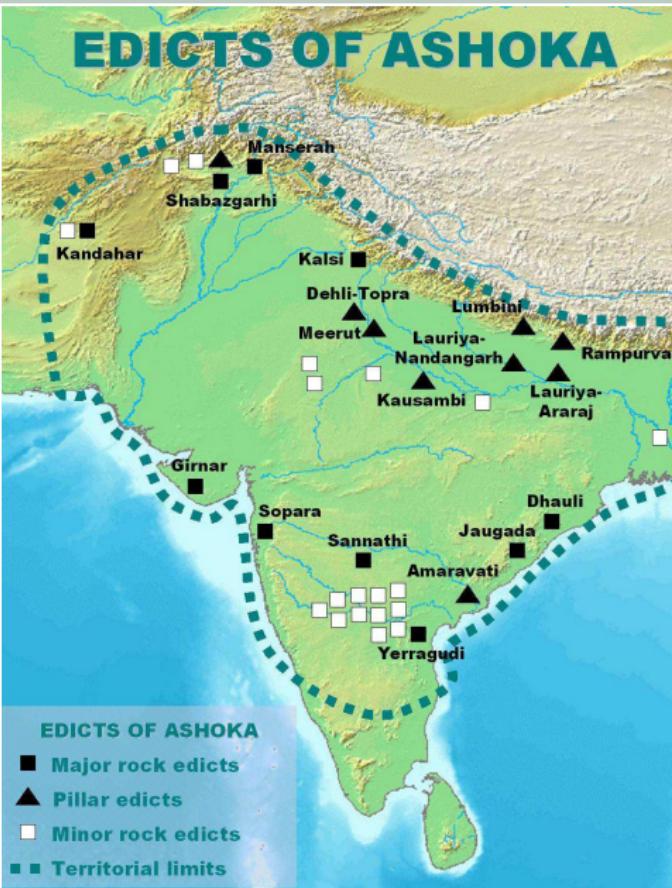
Pillars of Aśoka



Pillar of Vaishali

Pillars of Aśoka

- Aśoka Maurya: one of India's greatest emperors, 304–232 BC
 - *Sanskrit*: अशोक / अशोक + *aśoka* “without sorrow”
- Allegedly converted to Buddhism in remorse after a particularly bloody conquest
- Erected numerous pillars throughout his empire
 - 12–15m (40–50') high
 - Inscribed with various edicts about morality, based on Buddhist beliefs
 - Erected at Buddhist monasteries, important sites in the Buddha's life, places of pilgrimage
 - Only 19 inscribed columns have survived
- Buddhism spread rapidly with Aśoka's support



Roman alphabet

ooooooooooooooo

N'Ko

ooooooo

Ge'ez

oooooooooooooo

Indic Scripts

oooooooooooooooooooo●oooooooo

Devanāgarī

oooooooooooooooooooo



Lion Capital of Aśoka, from the Pillar of Sarnath



सत्यमेव जयते

National Emblem of India.
⟨satyameva jayate⟩ “Truth alone triumphs”



Carving of a *chakravartin* “universal ruler”, possibly Aśoka, with *dharmachakra* “wheel of dharma”. 1st century BC.



Flag of India, with *dharmachakra*

Aśoka's First Pillar Edict, ~242 BC

न दातुर्वा उमाहू ज्ञेष्ठाक्षे न द्विग्नाहृष्टालूः
 द्वैत्यु चैत्यु उपायु ज्ञात्यु क्षयु अलू द्वैत्यु वैत्यु
 क्षत्यु उद्वैत्यु द्वैत्यु द्वैत्यु प्रवैत्यु प्रवैत्यु प्रवैत्यु
 लैत्यु प्रवैत्यु प्रवैत्यु क्षत्यु उद्वैत्यु क्षत्यु प्रवैत्यु
 लैत्यु क्षत्यु उद्वैत्यु द्वैत्यु द्वैत्यु द्वैत्यु द्वैत्यु
 द्वैत्यु द्वैत्यु

(devānampiye piyadasi läja hevaṇ āha sađuvīsativasābhisisitena me iyam
 dhammalipi likhāpita hidatapālate dusam̄paṭipādaye am̄nata agāya dhammakāmatāya
 anusathiya dhammāpekha dhammakāmatā ca suve suve vadhitā vadhisati ceva pulisā pi me
 ukasā ca gevayā ca majhimā ca anuvidhīyam̄ti sampaṭipādayam̄ti ca alam̄ capalam̄ samādāpayitave
 hemeva am̄tamahāmātā pi esā hi vidhi yā iyam dhammena pālana dhammena vidhāne dhammena sukhīyana
 dhammena gotī ti)

Beloved-of-the-Gods, King Priyadarśi, speaks thus: This Dharma edict was written twenty-six years after my coronation. Happiness in this world and the next is difficult to obtain without much love for the Dharma, much self-examination, much respect, much fear (of evil), and much enthusiasm. But through my instruction this regard for Dharma and love of Dharma has grown day by day, and will continue to grow. And my officers of high, low, and middle rank are practicing and conforming to Dharma, and are capable of inspiring others to do the same. Mahamatras in border areas are doing the same. And these are my instructions: to protect with Dharma, to make happiness through Dharma and to guard with Dharma.

Aśoka's First Rock Edict, ~242 BC

:: श.०४८८६३९४८९४७१
 एवं लिखा तथा निर्माण कर्ता प्रियदर्श
 एवं देवता एवं विश्वामित्र
 एवं अमृता एवं अमृता एवं अमृता
 एवं अमृता एवं अमृता एवं अमृता

Beloved-of-the-Gods, King Priyadarśī, has caused this Dhamma edict to be written. Here (in my domain) no living beings are to be slaughtered or offered in sacrifice. Nor should festivals be held, for Beloved-of-the-Gods, King Priyadarśī, sees much to object to in such festivals, although there are some festivals that Beloved-of-the-Gods, King Priyadarśī, does approve of.

Formerly, in the kitchen of Beloved-of-the-Gods, King Priyadarśī, hundreds of thousands of animals were killed every day to make curry. But now with the writing of this Dhamma edict only three creatures, two peacocks and a deer are killed, and the deer not always. And in time, not even these three creatures will be killed.

Sanskrit grammarians

- Oral tradition of linguistic analysis, later written down
- Pāṇini: early Indian linguist, 4th century BC
- Wrote one of the earliest grammars of Sanskrit
 - No longer spoken natively by this point!
- Order of graphemes based on this analysis
- Arranged by place of articulation, back to front:
 - Velar, palatal, retroflex, dental, labial
 - Further divided by voice features/manner features

ooooooooooooooo

ooooooo

oooooooooooooo

oooooooooooooooooooo●oooo

ooooooooooooooo



A copy of Pāṇini's Sanskrit grammar

Origins of the scripts

- (These are described at length in your textbook!)
- Kharoṣṭhī
 - Most likely derived from Aramaic abjad
 - Similar shapes of many graphemes
- Brāhmī: several possibilities
 - An original invention
 - Not substantiated
 - Still a large originally Indian component
 - Restructured based on analysis by Indian grammarians
 - Derived from Indus script
 - Not likely
 - Indus society collapsed before arrival of Indo-Aryans

Origins of the scripts

- Based on a Semitic abjad
 - Plausible, but evidence not as clear as for Kharoṣṭhī
- Based on Kharoṣṭhī
 - Most plausible
 - Nearly identical structure
 - Differences possibly result of conscious attempt to make a different-looking script

Initial Vowels

- What happens if a word begins with a vowel?
 - In Ge'ez, this did not happen: all syllables were CV
 - In other Ethiopian Semitic languages, vowel-initial syllables are allowed
 - For those languages, the /?/ form is used, even if the initial glottal stop isn't pronounced

/ə/ /u/ /i/ /a/ /e/ /ɪ/ /o/
 /?V/ or /#V/       

- In Indic abugidas, there is often a *distinct* initial free allograph for vowels, where there is no free consonant grapheme for the bound allograph to attach to

Vowels in abudigas

- Kharosthī and Brāhmī are abugidas that allow vowel-initial words
 - Consonant graphemes have **inherent vowel** (often / a /)
 - Other vowels added as diacritics to consonant graphemes
 - Special strategy needed for stand-alone vowels
- Kharosthī

- Same diacritics used for initial vowels, bound to free ⟨a⟩

/ a /	/ i /	/ u /	/ r̥ /	/ e /	/ o /
/ V /	'	σ	s	'	r̥
/ k /	ঢ	ঢ	ঢ	ঢ	ঢ
/ #V /	ঢ	ঢ	ঢ	ঢ	ঢ

Brāhmī vowel diacritics

- Brāhmī represented vowel length
- Distinct free graphemes for each initial vowel!
 - This strategy is adopted by most of the Brahmic daughter languages

	a	i	u	r̥	l̥	e	o
	ā	ī	ū	ṛ̥	ṝ̥	ai	au
V	᳚	᳔	᳖	᳚	᳗	᳚	᳖
	᳚	᳔	᳖	᳚	᳗	᳚	᳖
k	+	+	+	+	+	+	+
	+	+	+	+	?	+	+

Devanāgarī

Devanāgarī

- Major writing system in India and Nepal
- The most widespread of the Brahmic scripts, one of the most widely-used scripts in the world
- Used for over 120 languages
 - Hindi, Marathi, Nepali, ...

हर्षोऽननिसमदिविजं ददानेदमहत्तदिकं। अग्रगामाधाराहृवेति येवानामेऽक्षणं देवास्तरागदिप्रतिक्षेप्ति
 जिवेधने॥ आश्रमस्यौकृतिः काप्रकृष्णस्मृतिः॥ एजामिषेकं प्रयमस्मृतेकर्तव्यानिवाच॥ अनप्राप्तन
 मारुपोगसा लांचप्रवेशाने॥ सूर्यनविद्यासौनामानिपन्नं देवमरोम्बवः द्रवतारभस्त्रातीवक्षाम्यकामं दया
 प्रत्यांक्षप्रवित्तेनुरुद्धरेष्मलभासेद्विवर्जयेत्॥ प्रसादानन्विः कामापि काम्यएतत्क्षेत्रे विवेधान्॥ अ
 ग्रामधारप्रसर्वेषाप्यद्विताधान स्तैसिलिकल्पेनकर्तव्यानाम्॥ अक्षरायशुलेसानुपर्याविद्विसे
 लेनक्षिललंतरंसंभवान्॥ तथ्यद्यमठक्षयं॥ अलेनमग्रेत्कुर्मन्त्रिसामेनिकोऽद्विपाणा
 द्विकर्त्तिविसामविमतिक्षुवेष्वह्नि स्यायपर्याधानवानुप्रीसादिकमध्यिभासालयाणका आ
 द्योपाकर्त्तिविकर्त्तयन्॥ स्मृष्टप्राप्तविशेषात्याविद्वितंवर्जयेभले इति॥ ननुसेतवागाहैसलसाप्तवि
 षेष्वनपुरुषं॥ कुतिविहृतवसंतादैः॥ सूर्यसांकोषादीपादितिवेत्॥ त्रिन्यानर्थकोषप्राणानांविप
 द्वितंवलावलत्तिमायाद्वन्नुत्तिः॥ सावकाशः स्मृतिश्वविरवकाशान्त्रस्वसामिक्षुनिवाधान्॥ अ
 तदेवक्षाधवविहानेश्वरपतितवज्ज्ञानविष्पत्तेनपावकानायामाहितानिमित्तिर्हं
 तियत्तपात्रैस्त्रेतिक्षुतेः पनितामादिप्रतिपत्तिविधाविक्षयकैताम्यप्रस्तुत्तुमुक्तावत्येवत्तुम्

Devanāgarī vowel diacritics

- Some vowels come linearly *before* their consonant!
- This is a feature of many Brahmic writing systems

a	i	u	ṛ	ṝ	e	o
ā	ī	ū	ṝ	ṝ	ai	au
V	अ	इ	उ	ऋ	ए	ओ
	आ	ई	ऊ		ऐ	औ
k	क	कि	कु	कृ	के	को
	का	की	कू	कृ	कै	कौ

Virāma

- Diacritic called **virāma** cancels the inherent vowel
 - Like the *sukun* we saw in Arabic

प /pa/ प् /p/

त /ta/ त् /t/

क /ka/ क् /k/

कप /kapa/

कप् /kap/

Inherent Vowels

- Basic graphemes with inherent vowel: core property of abugidas
- *Pronunciation* of inherent vowel can vary between languages
- The inherent vowel is often but **not always** pronounced /a/
 - (Ethio-Semitic: /ə/)

	Hindi	Bengali
प	/pa/	প /pɔ/
ত	/ta/	ত /tɔ/
ট	/ʈa/	ট /ʈɔ/
ক	/ka/	ক /kɔ/

Conjuncts

- Ligatures of several graphemes, used for clusters of consonants
- Only the last consonant in the cluster is followed by a vowel
- Usually conjuncts are formed from just two consonants

ग + य = ग्य
 /ga/ + /ja/ = /gja/

- Sometimes larger conjuncts are possible!

द + ध + र + य =
 /da/ + /dʱa/ + /ra/ + /ja/



/ddʱrja/

Conjunct formation

- Most Devanāgarī consonants lose a vertical stem when they are the first element in a conjunct
 - The stemless half form is a **bound allograph**

ण + ट = न्ट
 /ɳa/ /ʈa/ /ɳʈa/

	/ k ^h a /	/ ga /	/ g ^h a /	/ tʃa /	/ dʒa /	/ dʒ ^h a /	/ pa /
free	ख	ग	घ	च	ज	झ	ञ
bound	खु	गु	टु	चु	जु	झु	ञु
	/ -g ^h a /	-घ	ख्घ	ग्घ	छ्घ	च्घ	ज्घ

- Graphemes without a vertical stem reduce in different ways

Conjuncts

- Formation of conjuncts is not always transparent!
- Forms can be slightly irregular, or totally unpredictable

क + य = क्य
 /ka/ + /ja/ = /kja/

क + ष = क्ष
 /ka/ + /ṣa/ = /kṣa/

Conjunct formation

- The grapheme ⟨r⟩ has **two bound allographs**, depending on whether it is the first element in a conjunct
 - *First in conjunct*—curve above stem
 - *Not first in conjunct*—stroke at the bottom

र	+	क / ka /	→	कं / rka /
र / ra /	+	ख / kʰa /	→	खं / rkʰa /
र	+	ग / ga /	→	गं / rga /
र	+	ঘ / gʱa /	→	ঘং / rgʱa /
ক / ka /	+	র	→	ক্ / kra /
খ / kʰa /	+	র / ra /	→	খ্ / kʰra /
গ / ga /	+	র	→	গ্ / gra /
ঘ / gʱa /	+	র	→	ঘ্ / gʱra /

Conjunct formation (bonus!)

- Some graphemes have a stacked bound allograph, which appears below the free grapheme
 - This is only seen in geminates
 - Not totally predictable by shape: ଙ is a ‘stem’ grapheme

ଟ / t̪a /	+	ଟ / t̪a /	→	ଟ୍ଟ	/ t̪t̪a /
ଟ / t̪a /	+	ଠ / t̪ha /	→	ଟ୍ଟଠ	/ t̪t̪ha /
ଠ / t̪ha /	+	ଠ / t̪ha /	→	ଠ୍ଠ	/ t̪t̪ha /
ଡ / d̪a /	+	ଡ / d̪a /	→	ଡ୍ଟ	/ d̪d̪a /
ଡ / d̪a /	+	ଢ / d̪ha /	→	ଡ୍ଟଢ	/ d̪d̪ha /
ଢ / d̪ha /	+	ଢ / d̪ha /	→	ଢ୍ଢ	/ d̪d̪ha /
ଙ / l̪a /	+	ଙ / l̪a /	→	ଙ୍ଙ	/ l̪l̪a /

Conjunct formation (bonus!)

- Conjuncts beginning with / d / or / h / have the second consonant fused below or to the right

द	+	द / da /	→	द्द / dda /
द	+	ध / dʱa /	→	द्ध / ddʱa /
द	+	ब / ba /	→	द्ब / dba /
द / da /	+	भ / bʱa /	→	द्भ / dbʱa /
द	+	व / va /	→	द्व / dva /
द	+	म / ma /	→	द्म / dma /
द	+	य / ya /	→	द्य / dyā /
<hr/>				
ह	+	म / ma /	→	ह्म / hma /
ह / ha /	+	य / ya /	→	ह्य / hya /

Conjunct formation (bonus!)

- However, there are still exceptions...
- Some have decomposable shapes, but others are totally opaque

त / ta /	+	र / ra /	→	त्र / tra /
श / ūa /	+	र / ra /	→	श्र / ūra /
त / ta /	+	त / ta /	→	त्त / tta /
क / ka /	+	त / ta /	→	क्त / kta /
ज / dʒa /	+	ञ / ūna /	→	ञ्ज / dʒūna / [gja]
क / ka /	+	ष / ūsa /	→	क्ष / kṣa /

Conjunct formation (bonus!)

- Let's revisit that monster conjunct now...

द + ध + र + य =
 / da / / d^ha / / ra / / ja /



/ dd^hrja /

- We've seen how / d / and / d^h / combine (झ), how / r / attaches below (ڑ), and how these then fuse with / ja / (य)

Conjuncts vs. virāma

- Not all consonant combinations actually appear as conjuncts!
- Conjuncts are usually used in loanwords from Sanskrit or Pāli
- The virāma diacritic is commonly used to indicate lack of inherent vowel in modern Hindi words

Inherent vowels & orthographic depth

- However, not all clusters and coda consonants are orthographically indicated
- Some codas and clusters result from historical deletion of default /a/

Sanskrit /ratʃana:/ रचना ⟨ratʃana:⟩

Hindi /ratʃna:/ रचना ⟨ratʃana:⟩ not रच्ना

Sanskrit /na:ma/ नाम ⟨na:ma⟩

Hindi /na:m/ नाम ⟨na:ma⟩ not नाम्

Vowels & orthographic depth

- English has had similar phonological changes, historically
- The writing system was not updated!
 - Phonological changes without orthographic changes → orthography becomes deeper

Middle English /na:mə/ ⟨name⟩

Modern English /nejm/ ⟨name⟩ *⟨nam⟩

Roman alphabet

ooooooooooooooo

N'Ko

ooooooo

Ge'ez

oooooooooooo

Indic Scripts

oooooooooooooooooooooooooooo

Devanāgarī

oooooooooooooooo

Transliteration

पैजामा
खकि
योग

Transliteration

पैजामा

⟨pajāmā⟩ ‘pyjamas’

योग

⟨yoga⟩ ‘yoga’

खकि

⟨khaki⟩ ‘khaki’