

# Writing Systems

Zdeněk Žabokrtský

partially based on slides of Magda Ševčíková and Anna Nedoluzhko

## Warm-up exercise - try to align languages and scripts

- |             |                 |
|-------------|-----------------|
| (1) 𑌵𑍇𑍄𑍅    | A. Chinese      |
| (2) λογική  | B. Arabic       |
| (3) ተፈጽኗል   | C. Hebrew       |
| (4) 𑌨𑌸𑌹𑌺    | D. Armenian     |
| (5) nguòl   | E. Neo-Tifinagh |
| (6) ひらがな    | F. Greek        |
| (7) 𐤀𐤂𐤁𐤅    | G. Latin        |
| (8) 인간은     | H. Cyrillic     |
| (9) ყველა   | I. Ethiopic     |
| (10) 𑌮𑌰𑌭    | J. Georgian     |
| (11) 𑌮𑌰𑌭𑌰𑌭  | K. Thai         |
| (12) 漢      | L. Hiragana     |
| (13) ДОЛЖНЫ | M. Devanagari   |
| (14) մարդիկ | N. Mongolian    |
| (15) الناس  | O. Hangul       |

# Outline

1. Intro, basic notions, writing vs. speech
2. Historical development of writing systems
3. Selected writing systems used today
4. A global view on world's writing systems
5. Final remarks

# Introduction

# Writing

writing - a system of recording language by means of visible or tactile marks

- the communication of meaning is the primary purpose of most writing
- today, more communication takes place in the written than in the oral form
- writing is not a “minor” form of communication

# Writing vs. speech

- discrete
  - timeless
  - autonomous
  - permanent
  - visible
  - produced by hand
- continuous
  - bound to utterance time
  - contextual
  - temporal
  - audible
  - produced by voice

# Writing systems of the world

- the first language technology!
- enormous diversity
- writing systems virtually located on a scale between semiography and phonography
  - semiography – meaning-based writing
  - phonography – sound-based writing
  - ideal, but not realizable writing systems
- no writing system is immediately comprehensible without instructions, initiation needed
- quite different learning curves/cognitive loads though

# Five basic types of modern writing systems

1. **alphabets** – a symbol corresponds to a sound (a consonant or a vowel)
  - example: Latin(=Roman)/Cyrilic/Greek scripts
2. **abjads** – a symbol corresponds to a consonant (vowels to be inferred by readers)
  - example: Arabic and Hebrew scripts
3. **abugidas** – a symbol corresponds to a consonant-vowel pairing (vowels rendered as modifications of the base consonant shape)
  - example: Devanagari (used for Hindi)
4. **syllabaries** – a symbol corresponds to a syllable or a mora
  - example: Japanese Kana
5. **logo-syllabaries** – a symbol corresponds to a word or a morpheme
  - example: Chinese characters



# Other basic notions

- a script = a set of symbol/character/sign/letter/glyph/marks
- an orthography = rules (conventions) for spelling, punctuation, capitalization...
- a writing system = a script + an orthography

Examples:

- the English writing system → alphabetic, based on the Latin script, with English spelling rules.
- the Japanese writing system → a combination of logographic script (kanji) + syllabic scripts (hiragana & katakana)
- the Hindi writing system → an abugida using Devanagari with Hindi orthographic rules

# Historical development of writing system

# Proto-writing – visible communication marks

Cave paintings



Turtle plastrons



Clay slabs



# Invention of writing

- a relatively recent invention, if compared to (in years before present):
  - emergence of anatomically modern humans around 300 kYBP
  - complex tools around 150 kYBP
  - language-capable cognition around 100 kYBP
  - migration out of Africa around 60 kYBP
  - cultural explosion 40 kYBP
  - agriculture 10 kYBP
- writing invented independently (at least) four times:
  - **Mesopotamia** – 3400 BC
  - **Egypt** – 3200 BC
  - **China** – 1300 BC
  - **Mesoamerica** – 900 BC
- very gradual

































# Sumerian writing

- southern Mesopotamia, around 3100 BC
- clay tablets, heavy but indestructible
- used a reed stylus to press wedge-shaped marks
- mostly used for record-keeping (agriculture, trade, inventories)



# Sumerian writing, cont.

- early tables – the primary referent of the signs were physical objects (animals and their body parts, tools, plants...)
- pictograms developed into cuneiform signs; around 1000 signs
- original meaning: concrete referents, abstract meaning added later (star → God)
- phonetic value added too: a picture of hand, 'hand' in Sumerian: 'šu', a picture of hand = phonetic syllable 'šu'

	3200 BCE	3000 BCE	2400 BCE	1000 BCE
sag 'head'				
gin 'to walk'				
šu 'hand'				
še 'barley'				
ninda 'bread'				
a 'water'				
ud 'day'				
mušen 'bird'				

# Egyptian hieroglyphs

- probably developed from preliterate visual art (proto-hieroglyphic paintings 3500 BC...)
- 900 signs in the Middle Kingdom period (around 2000 BC)
- combined logographic, syllabic and alphabetic elements (e.g. phonetic glyphs representing consonants)



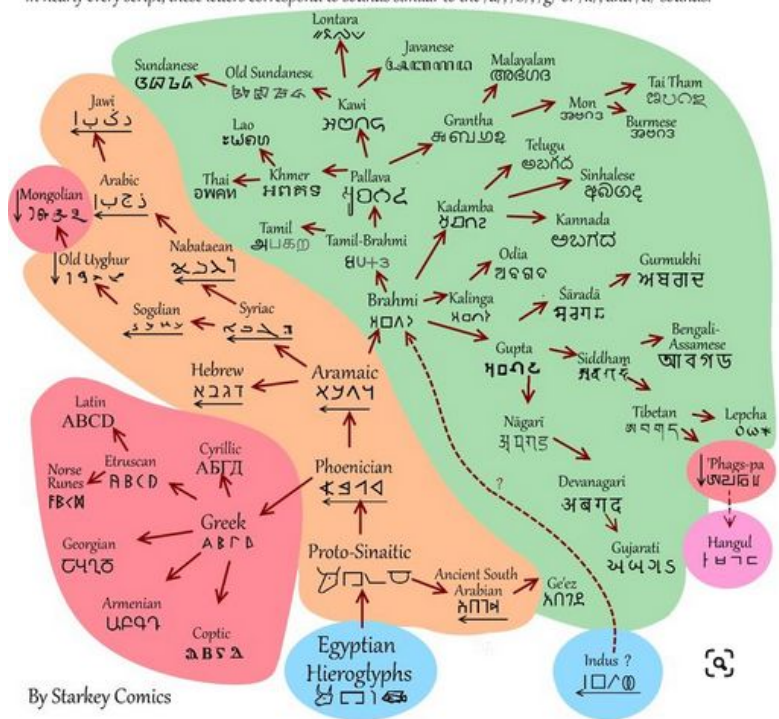


# From hieroglyphs to today's writing system

**The ABCD Family Tree**

True Alphabets    Abjads  
Featural Alphabets    Abugidas  
Logograms

This infographic shows how 57 different scripts are descended from ancient Egyptian hieroglyphs. The colours show the type of script, and the red arrows show how the scripts are related. For scripts which are not read from left to right, the black arrows show the direction of writing. Where possible, letters corresponding to the A, B, C, and D symbols in the Latin alphabet are shown. In nearly every script, these letters correspond to sounds similar to the /a/, /b/, /g/ or /k/, and /d/ sounds.



By Starkey Comics

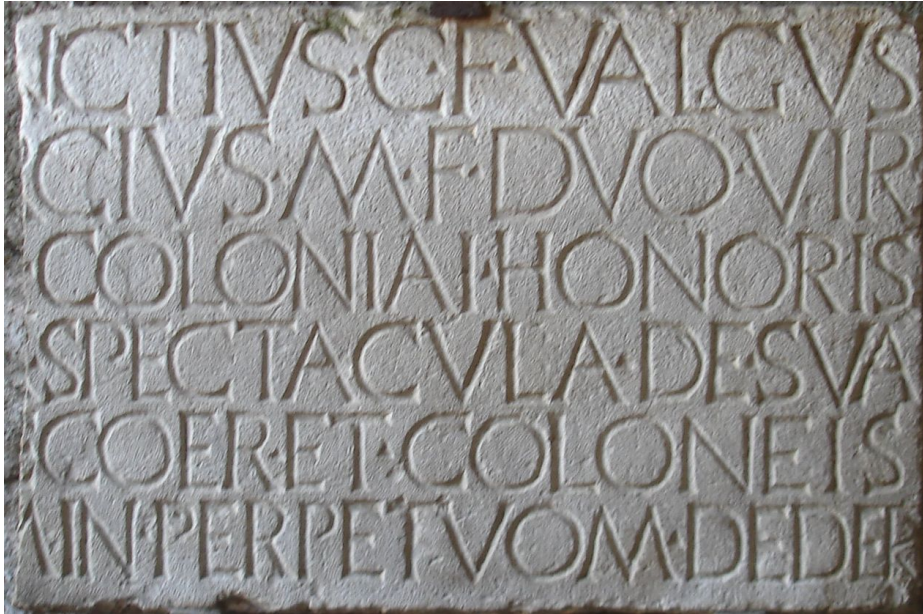
- Exercise: could you align the trajectory from the Egyptian Hieroglyphs to e.g. the Czech alphabet with the time axis?



## **Selected writing systems used today**

# Latin (Roman) alphabet

a Pompeii amphitheatre inscription



# Latin alphabet, cont.

- various modifications:
  - sometimes more consonants need to be distinguished:
    - Wykład o systemach pisma
  - or more vowels:
    - Előadás az írásrendszerekről
  - or both:
    - Ein Vortrag über **S**chriftsysteme
    - Přednáška o pís**m**ových syst**é**mech
  - or tones:
    - Bài giảng về hệ thống chữ viết
  - or clicks:
    - !Kung (a language spoken in Namibia)

# Latin alphabet, cont.

- not a perfect (one-to-one) alphabet
  - a single symbol pronounced differently in English: cat, city, cello, ocean, muscle
  - many non-one-to-one exceptions even in relatively phonetic orthography of Czech (i/y, ú/ů, mně/mě, s/z, d/t ...)
  - a crosslingual perspective: (almost) the same sound written differently across languages – č cz tsch cs ch various modifications:

- a sample from the Universal Declaration of Human Rights (the 1st article)

सभी मनुष्यों को गौरव और अधिकारों के मामले में जन्मजात स्वतन्त्रता और समानता प्राप्त है। उन्हें बुद्धि और अन्तरात्मा की देन प्राप्त है और परस्पर उन्हें भाईचारे के भाव से बर्ताव करना चाहिए।

## Devanagari, cont.

- a left-to-right abugida
- based on the ancient Brahmi script,
- used in a number of languages the northern Indian subcontinent, such as Hindi, Marathi, Kashmiri, Nepali... but not used in Dravidian languages
- round shapes, a horizontal line connecting letters
- 48 primary characters, including 14 vowels and 34 consonants

# Devanagari, cont. – an example of vowels in Hindi

अ	आ	इ	ई	उ	ऊ	ए
a	ā	i	ī	u	ū	e
[ʌ]	[a]	[i]	[i:]	[u]	[u:]	[e]
प	पा	पि	पी	पु	पू	पे
pa	pā	pi	pī	pu	pū	pe
ऐ	ओ	औ	अं	अः	अँ	ऋ
ai	o	au	aṁ	aḥ	āṁ	r
[æ:]	[o]	[ɔ:]	[aŋ]	[əh]	[ā:]	[ɾ]
पै	पो	पौ	पं	पः	पाँ	पृ
pai	po	pau		paḥ	pāṁ	pr̥

- a sample from UHDR in Korean

모든 인간은 태어날 때부터 자유로우며 그 존엄과 권리에 있어 동등하다. 인간은 천부적으로 이성과 양심을 부여받았으며 서로 형제애의 정신으로 행동하여야 한다.



# Hangul, cont.

- used for Korean (both in South and North Korea)
- created in 1443 CE by King Sejong the Great, letters for consonants reflect the shape of the speech organs
- modern Hangul: 14 consonant letters and 10 vowel letters
- technically an alphabet (one symbol – one sound)...
- ...however, with letters written in morphosyllabic blocks (at least one consonant and one vowel), not separately
- historically written top-to-bottom, nowadays left-to-right
- in a sense, one of the most scientific writing systems in the world
- also called a featural writing system – type of writing in which the shapes of the symbols systematically the phonetic features of the sounds they encode. In other words, the way a letter looks tells you something about how it is pronounced.

Korean Alphabet Chart

Consonants	Vowels									
	ㅏ (a)	ㅑ (ya)	ㅓ (o)	ㅕ (yo)	ㅗ (oh)	ㅛ (yo)	ㅜ (ow)	ㅠ (you)	ㅡ (er)	ㅣ (ee)
ㄱ(G)	가	가	거	겨	고	교	구	규	그	기
ㄴ(N)	나	냐	너	녀	노	뇨	누	뉴	느	니
ㄷ(D)	다	다	더	더	도	도	두	두	드	디
ㄹ(R/L)	라	랴	러	려	로	료	루	류	르	리
ㅁ(M)	마	마	머	며	모	묘	무	뮤	므	미
ㅂ(B)	바	바	버	벼	보	보	부	뷰	브	비
ㅅ(S)	사	샤	서	셔	소	쇼	수	슈	스	시
ㅇ Silent	아	야	어	여	오	요	우	유	으	이
ㅈ(J)	자	자	저	져	조	조	주	주	즈	지
ㅊ(CH)	차	차	쳐	쳐	초	초	추	추	츠	치
ㅋ(K)	카	카	커	켜	코	코	쿠	큐	크	키
ㅌ(T)	타	타	터	터	토	토	투	투	트	티
ㅍ(P)	파	파	퍼	펴	포	포	푸	퓨	프	피
ㅎ(H)	하	하	허	혀	호	호	후	휴	흐	히

©Modern Seoul Magazine (2012)

- a sample from UHDR in Arabic

## المادة 1

يولد جميع الناس أحرارًا متساوين في الكرامة والحقوق. وقد وهبوا عقلاً وضميرًا وعليهم أن يعامل بعضهم بعضًا بروح الإخاء.

# Arabic script, cont.

- used as the primary script for various language families because of the Quran
- an abjad for the Arabic language (letters only for consonants, vowels can be indicated by vowel marks)
- fits well the inflection nature of Arabic (k-t-b)
- some modifications for other languages used as true alphabets with mandatory vowels, e.g. for Kashmiri or Persian (the Perso-Arabic script)
- no capital letters

## Arabic script, cont.

ا	ب	ت	ث	ج	ح	خ
'a	b	t	t	ġ	h	ḥ
د	ذ	ر	ز	س	ش	ص
d	ḏ	r	z	s	š	ṣ
ض	ط	ظ	ع	غ	ف	ق
ḍ	t	ẓ	'	ġ	f	q
ك	ل	م	ن	ه	و	ي
k	l	m	n	h	w	y

人人生而自由，在尊严和权利上一律平等。他们赋有理性和良心，并应以兄弟关系的精神相对待。

# Chinese writing system

- traditional and simplified forms
- going back to 1400 BC (scripts on “oracle bones”)
- oldest writing system used continuously
- usually 1 character = 1 morpheme
- square-shaped, made of strokes
- over 100,000 characters

人類社会のすべての構成員の固有の尊厳と平等で譲ることのできない権利とを承認することは、世界における自由、正義及び平和の基礎であるので、



# Japanese writing system, cont.

- The main components:
  - Kanji (漢字)
    - thousands of characters, each with one or more readings
    - represent meaning and sometimes sound
    - used for nouns, verbs, adjectives, and roots of words
  - Hiragana (ひらがな)
    - syllabary for native Japanese words and grammatical elements
    - 46 basic symbols representing syllables (e.g., か = “ka”).
    - often used for inflections, particles, and words without kanji.
  - Katakana (カタカナ)
    - syllabary for foreign words, loanwords, onomatopoeia, and emphasis.
    - same sound system as hiragana but different symbols.

# Japanese writing system, cont.

- Hiragana

N	w	r	y	m	h	n	t	s	k	
ん N	わ WA	ら RA	や YA	ま MA	は HA	な NA	た TA	さ SA	か KA	a
		り RI		み MI	ひ HI	に NI	ち CHI	し SHI	き KI	i
		る RU	ゆ YU	む MU	ふ FU	ぬ NU	つ TSU	す SU	く KU	u
		れ RE		め ME	へ HE	ね NE	て TE	せ SE	け KE	e
	を WO	ろ RO	よ YO	も MO	ほ HO	の NO	と TO	そ SO	こ KO	o

- Katakana

n	w	r	y	m	h	n	t	s	k	
ン N	ワ WA	ラ RA	ヤ YA	マ MA	ハ HA	ナ NA	タ TA	サ SA	カ KA	a
	キ KI	リ RI		ミ MI	ヒ HI	ニ NI	チ CHI	シ SHI	キ KI	i
				ム MU	フ FU	ヌ NU	ツ TSU	ス SU	ク KU	u
				モ MO	ヘ HE	ネ NE	テ TE	セ SE	ケ KE	e
	ヲ WA	ロ RO	ヨ YO	モ MO	ホ HO	ノ NO	ト TO	ソ SO	コ KO	o

- Examples of Kanji

数 Number, Count	多 many, much	少 a few, a little	万 10,000	半 half	形 shape	太 thick	細 thin	広 wide	長 long	点 point	丸 circle	交 intersect	光 light	角 corner	計 measure	直 straight	線 line
矢 arrow	弱 weak	強 strong	高 high	同 same	親 parent	母 mother	父 father	姉 older sister	兄 older brother	弟 younger brother	妹 younger sister	自 oneself	友 friend	体 body	毛 hair	頭 head	顔 face
首 neck	心 heart	時 time	曜 weekday	朝 morning	昼 daytime	夜 night	分 minute, understand	週 week	春 spring	夏 summer	秋 autumn	冬 winter	今 now	新 new	古 old	間 interval	方 direction
北 north	南 south	東 east	西 west	遠 far	近 near	前 before	後 after	内 inside	外 outside	場 place	地 ground	国 country	園 garden	谷 valley	野 field	原 meadow	里 hometown
市 city	京 capital	風 wind	雪 snow	雲 cloud	池 pond	海 sea	岩 rock	星 star	室 room	戸 door	家 house	寺 Buddhist temple	通 pass through	門 gate	道 road	話 talk	言 say
答 answer	声 voice	聞 hear	語 language	読 read	書 write	記 record	紙 paper	画 brush stroke	絵 picture	図 drawing	工 craft	教 teach	晴 clear	思 think	考 consider	知 know	才 ability
理 reason	算 calculate	作 make	元 origin	食 eat	肉 meat	馬 horse	牛 cow	魚 fish	鳥 bird	羽 feather	鳴 chip	麦 wheat	米 rice	茶 tea	色 colour	黄 yellow	黒 black
来 come	行 go	帰 return	歩 walk	走 run	止 stop	活 active	店 store	買 buy	売 sell	午 noon	汽 steam	弓 bow (noun)	回 times	会 meeting	組 team	船 ship	明 bright
社 company	切 cut	電 electricity	毎 every	合 fit	当 thus, hit	台 pedestal	楽 pleasure	公 public	引 pull	科 section	歌 song	刀 sword	番 number	用 use	何 what	丁 pair	世 generation
両 both	主 master	乗 ride	予 beforehand	事 abstract thing	仕 serve	他 other	代 substitute	住 dwell	使 use	係 leader	倍 double	全 whole	具 tool	写 copy	列 row	助 help	勉 diligence

Find more at [palinu.deviantart.com](https://www.palinu.deviantart.com) © Basic 165, Revision: 1/3 (2020/08/08) Kanji Grade 2 + 3 (81 - 240) + (241 - 260) - Wallpaper Sheet made by Palinu based on 学年別漢字配当表

# **A global view on world's writing systems**

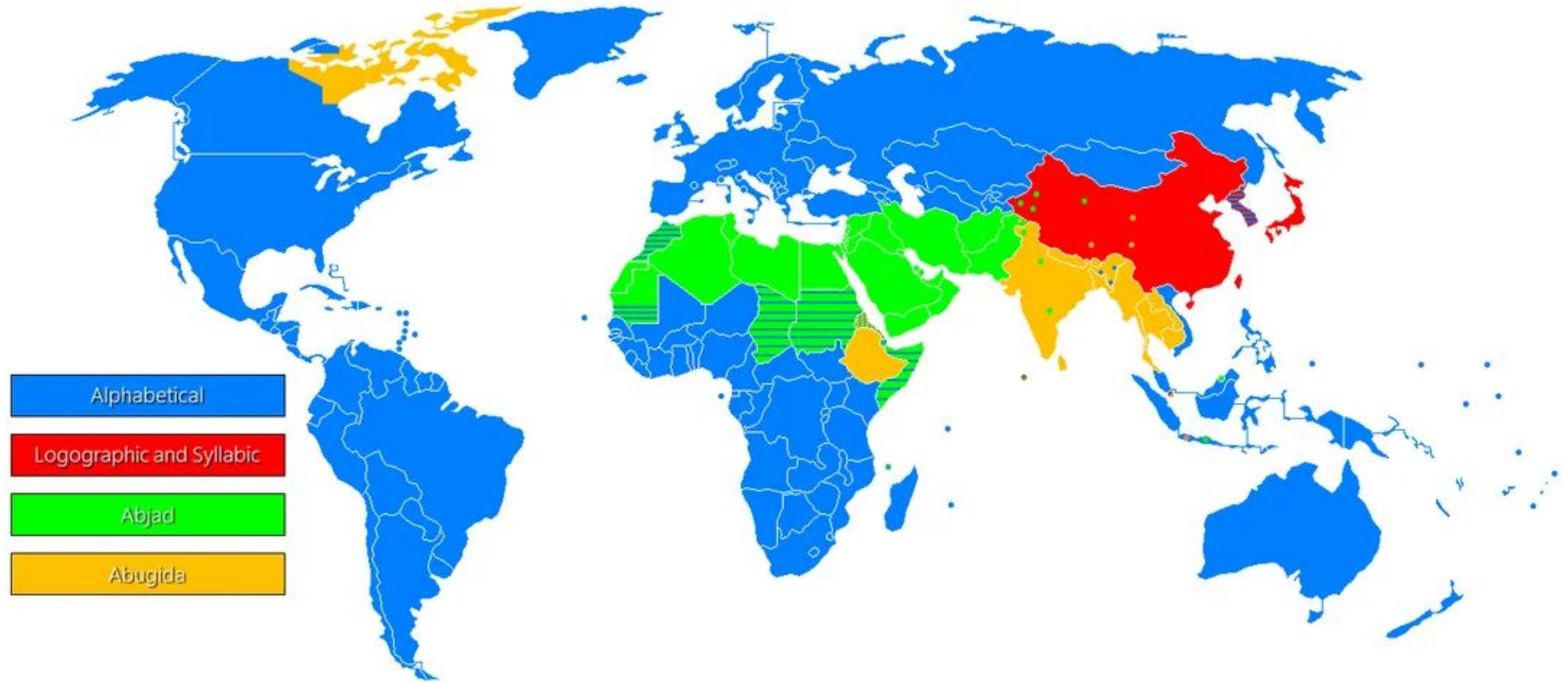
# The world's most popular scripts (worldatlas.com)

Rank	Name of script	Type	Population actively using (in millions)
1	Latin Latin	Alphabet	over 4900
2	Chinese 汉字汉字	Logographic	1340
3	Arabic العربية	Abjad	660+
4	Devanagari देवनागरी	Abugida	608+
5	Bengali-Assamese বাংলা-অসমীয়া	Abugida	300
6	Cyrillic Кириллица	Alphabet	250
7	Kana カナカナ	Syllabary	120
8	Javanese	Abugida	80
9	Hangul 한글 조선글	Alphabet, featural	78.7
10	Telugu తెలుగు	Abugida	74
11	Tamil தமிழ்	Abugida	70
12	Gujarati ગુજરાતી	Abugida	48
13	Kannada ಕನ್ನಡ	Abugida	45
14	Burmese မြန်မာ	Abugida	39
15	Malayalam മലയാളം	Abugida	38

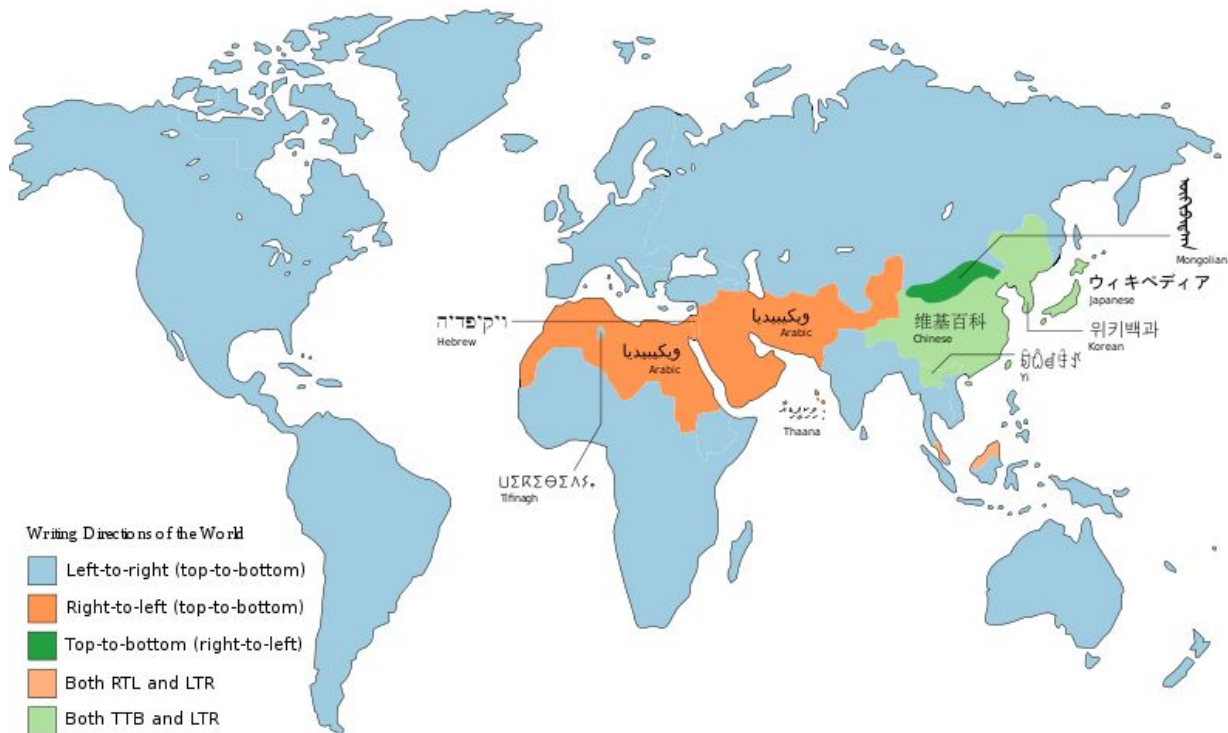
# The world's most popular scripts, cont.

16	Thai ไทย	Abugida	38
17	Sundanese	Abugida	38
18	Gurmukhi ਗੁਰਮੁਖੀ	Abugida	22
19	Lao ລາວ	Abugida	22
20	Odia ଓଡ଼ିଆ	Abugida	21
21	Ge'ez ግዕዝ	Abugida	18
22	Sinhala සිංහල	Abugida	14.4
23	Hebrew תיבתא	Abjad	14
24	Armenian Հայոց	Alphabet	12
25	Khmer ខ្មែរ	Abugida	11.4
26	Greek Ελληνικό	Alphabet	11
27	Lontara	Abugida	7.6
28	Tibetan བོད་	Abugida	5
29	Georgian ქართული	Alphabet	4.5
30	Modern Yī 𐆆𐆏	Syllabary	4
31	Mongolian ᠮᠣᠩᠭᠣᠯ	Alphabet	2
32	Tifinagh	Abjad	1
33	Syriac	Abjad	0.4
34	Thaana	Abugida	0.35
35	Inuktitut ᐃᓄᐅᐅᐅ	Abugida	0.035
36	Cherokee ᏍᏏᏉᏯ	Syllabary	0.02

# Type of the dominant writing system



# Writing direction



## **Final remarks**



# Take-home message

- large diversity (in spite of only a few roots)
- no writing system immediately intelligible
- writing systems invented independently a couple of times
- a scale from semiography to phonography (in reality always something in between)
- five basic types (vague boundaries): alphabets, abjads, abugidas, syllabaries, logo-syllabarie
- order-of-magnitude differences in sizes of character inventories

# Time for a discussion

- Relationships between a writing system and a language?
- Pros and cons of individual types of writing systems?
- Writing is not an object of linguistics (A paradox: linguistics tends to base the analysis on writing-induced concepts (e.g. a word) but subscribes to the principle of the primacy of speech)