Writing Systems

Zdeněk Žabokrtský
partially based on slides of Magda Ševčíková and Anna Nedoluzhko

January 4, 2024
A warm-up exercise

1. 
2. 
3. 
4. 
5. 
6. 
7. 
8. 
9. 
10. 
11. 
12. 
13. 
14. 
15. 

(a) Chinese
(b) Arabic
(c) Hebrew
(d) Armenian
(e) Neo-Tifinagh
(f) Greek
(g) Latin
(h) Cyrillic
(i) Ethiopic
(j) Georgian
(k) Thai
(l) Hiragana
(m) Devanagari
(n) Mongolian
(o) Hangul
Intro

Historical development of writing systems

Selected writing systems used today

A global view on world’s writing systems

Final remarks
Intro
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 (Internet)
writing is not a ‘minor’ form of communication
Writing vs. speech

- discrete
- continuous
- timeless
- bound to utterance time
- autonomous
- contextual
- permanent
- temporal
- visible
- audible
- produced by hand
- produced by voice
Writing systems of the world

- 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 instruction, 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
• 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
Historical development of writing systems
Proto-writing – visible communication marks

Things

Wampums (shell beads)

Pictograms (cave paintings)

Quipu from Inca people (talking knots)
Invention of writing

• a relatively recent invention
• invented independently (at least) four times:
  • Mesopotamia – 3400 BC
  • Egypt – 3200 BC
  • China – 1300 BC
  • Mesoamerica – 900 BC
• very gradual
• writing – the first language technology!
Sumerian writing

- southern Mesopotamia
- clay tables, heavy but indestructible
- economic records (more than 80% of the tablets)
Sumerian writing

• 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’
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 systems

The ABCD Family Tree

This infographic shows how 52 different scripts are descended from ancient Egyptian hieroglyphs. The colors 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/, /e/, /o/, /u/ and /i/ sounds.

Intro
Historical development of writing systems
Selected writing systems used today
A global view on world’s writing systems
Final remarks
Selected writing systems used today
Latin (Roman) alphabet

a Pompeii amphitheatre inscription
Latin alphabet

various modifications:

• sometimes more consonants need to be distinguished:
  • Wykład o systemach pisma
  • The lecture about writing systems

• or more vowels:
  • Előadás az írásrendszerekről

• or both:
  • Ein Vortrag über Schriftsysteme
  • Přednáška o písmových systémech

• or tones:
  • Bài giảng về hệ thống chữ viết

• or clicks:
  • Kung (a language spoken in Namibia)
not a perfect (one-to-one) alphabet

• a single symbol pronounced differently in English: cat, city, cello, ocean, muscle

• many non-1-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
a sample from UDHR (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 (Hindi, Marathi, Kashmiri, Nepali...)
- round shapes, a horizontal line connecting letters
- 48 primary characters, including 14 vowels and 34 consonants
Devanagari, cont. – an example of vowels in Hindi

\[\begin{array}{cccccccc}
\text{अ} & \text{आ} & \text{इ} & \text{ई} & \text{उ} & \text{ऊ} & \text{ए} \\
\text{a} & \text{a} & \text{i} & \text{iː} & \text{u} & \text{uː} & \text{e} \\
[\text{A}] & [\text{a}] & [\text{i}] & [\text{iː}] & [\text{u}] & [\text{uː}] & [\text{e}] \\
\text{प} & \text{पा} & \text{पि} & \text{पी} & \text{पु} & \text{पू} & \text{पे} \\
\text{pa} & \text{pā} & \text{pī} & \text{pī} & \text{pu} & \text{pu} & \text{pe} \\
[\text{p}] & [\text{pː}] & [\text{pː}] & [\text{pː}] & [\text{pː}] & [\text{pː}] & [\text{pː}] \\
\text{ऐ} & \text{ओ} & \text{औ} & \text{अ} & \text{ॐ} & \text{ऋ} \\
\text{ai} & \text{o} & \text{au} & \text{aː} & \text{aː} & \text{r} \\
[\text{æː}] & [\text{o}] & [\text{oː}] & [\text{aː}] & [\text{aː}] & [\text{r}] \\
\text{पै} & \text{पो} & \text{पाँ} & \text{पः} & \text{पः} & \text{पू} \\
\text{pai} & \text{po} & \text{pau} & \text{pā} & \text{pā} & \text{pu} \\
[\text{pː}] & [\text{pː}] & [\text{pː}] & [\text{pː}] & [\text{pː}] & [\text{pː}] \\
\end{array}\]
전체 인간은 태어날 때부터 자유로우며 그 존엄과 권리에 있어 동등하다. 인간은 천부적으로 이성과 양심을 부여받았으며 서로 형제애의 정신으로 행동하여야 한다.

a sample from UHDR in Korean
• 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
# Korean Alphabet Chart

<table>
<thead>
<tr>
<th>Consonants</th>
<th>Vowels</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(a)</td>
</tr>
<tr>
<td>ㄱ (G)</td>
<td>가</td>
</tr>
<tr>
<td>ㄴ (N)</td>
<td>나</td>
</tr>
<tr>
<td>ㄷ (D)</td>
<td>다</td>
</tr>
<tr>
<td>ㄹ (R/L)</td>
<td>라</td>
</tr>
<tr>
<td>ㅁ (M)</td>
<td>마</td>
</tr>
<tr>
<td>ㅂ (B)</td>
<td>바</td>
</tr>
<tr>
<td>ㅅ (S)</td>
<td>사</td>
</tr>
<tr>
<td>ㅇ Silent</td>
<td>아</td>
</tr>
<tr>
<td>ㅈ (J)</td>
<td>자</td>
</tr>
<tr>
<td>ㅊ (CH)</td>
<td>쳐</td>
</tr>
<tr>
<td>ㅋ (K)</td>
<td>카</td>
</tr>
<tr>
<td>ㅌ (T)</td>
<td>타</td>
</tr>
<tr>
<td>ㅍ (P)</td>
<td>파</td>
</tr>
<tr>
<td>ㅎ (H)</td>
<td>하</td>
</tr>
</tbody>
</table>

©Modern Seoul Magazine (2012)
Arabic script

a sample from UHDR in Arabic

المادة 1

يولد جميع الناس أحرارًا متساوين في الكرامة والحقوق. وقد وجبوا عفلاً وضميرًا
وعليهم أن يعامل بعضهم بعضاً بروح الأخاء.
• used as the primary script for various language families because of the Quaran
• an abjad for the Arabic language (letters only for consonants, vowels can be indicated by vowel marks)
  • fits well the introflection 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
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 characters (Hanzi)

- traditional and simplified forms
- going back to 1400 BC (scripts on “oracle bones”)
- oldest writing system used continuously
- usually 1 character = 1 word or morpheme
- square-shaped, made of strokes
- over 100,000 characters
A global view on world’s writing systems
The world’s most popular scripts (worldatlas.com)

<table>
<thead>
<tr>
<th>Rank</th>
<th>Name of script</th>
<th>Type</th>
<th>Population actively using (in millions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Latin Latin</td>
<td>Alphabet</td>
<td>over 4900</td>
</tr>
<tr>
<td>2</td>
<td>Chinese 漢字漢字</td>
<td>Logographic</td>
<td>1340</td>
</tr>
<tr>
<td>3</td>
<td>Arabic العربية</td>
<td>Abjad</td>
<td>660+</td>
</tr>
<tr>
<td>4</td>
<td>Devanagari देवनागरी</td>
<td>Abugida</td>
<td>608+</td>
</tr>
<tr>
<td>5</td>
<td>Bengali-Assamese বাংলা-অসমীয়া</td>
<td>Abugida</td>
<td>300</td>
</tr>
<tr>
<td>6</td>
<td>Cyrillic Кириллица</td>
<td>Alphabet</td>
<td>250</td>
</tr>
<tr>
<td>7</td>
<td>Kana かなカナ</td>
<td>Syllabary</td>
<td>120</td>
</tr>
<tr>
<td>8</td>
<td>Javanese</td>
<td>Abugida</td>
<td>80</td>
</tr>
<tr>
<td>9</td>
<td>Hangul 한글 조선글</td>
<td>Alphabet, featural</td>
<td>78.7</td>
</tr>
<tr>
<td>10</td>
<td>Telugu తెలుగు</td>
<td>Abugida</td>
<td>74</td>
</tr>
<tr>
<td>11</td>
<td>Tamil தமிழ்</td>
<td>Abugida</td>
<td>70</td>
</tr>
<tr>
<td>12</td>
<td>Gujarati ગુજરાતી</td>
<td>Abugida</td>
<td>48</td>
</tr>
<tr>
<td>13</td>
<td>Kannada ಕನ್ನಡ</td>
<td>Abugida</td>
<td>45</td>
</tr>
<tr>
<td>14</td>
<td>Burmese မြန်မာ</td>
<td>Abugida</td>
<td>39</td>
</tr>
<tr>
<td>15</td>
<td>Malayalam മലയാളം</td>
<td>Abugida</td>
<td>38</td>
</tr>
</tbody>
</table>
# The world’s most popular scripts (worldatlas.com), cont.

<table>
<thead>
<tr>
<th>No.</th>
<th>Script</th>
<th>Script Type</th>
<th>Popularity</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>Thai (แก)</td>
<td>Abugida</td>
<td>38</td>
</tr>
<tr>
<td>17</td>
<td>Sundanese</td>
<td>Abugida</td>
<td>38</td>
</tr>
<tr>
<td>18</td>
<td>Gurmukhi (ਘੁਰਮੁਖੀ)</td>
<td>Abugida</td>
<td>22</td>
</tr>
<tr>
<td>19</td>
<td>Lao (ລາວ)</td>
<td>Abugida</td>
<td>22</td>
</tr>
<tr>
<td>20</td>
<td>Odi (ଓଡ଼ିଆ)</td>
<td>Abugida</td>
<td>21</td>
</tr>
<tr>
<td>21</td>
<td>Greek (Γεωργιαν)</td>
<td>Abugida</td>
<td>18</td>
</tr>
<tr>
<td>22</td>
<td>Sinhala (සිංහල)</td>
<td>Abugida</td>
<td>14.4</td>
</tr>
<tr>
<td>23</td>
<td>Hebrew (עברית)</td>
<td>Abjad</td>
<td>14</td>
</tr>
<tr>
<td>24</td>
<td>Armenian (Հայերեն)</td>
<td>Alphabet</td>
<td>12</td>
</tr>
<tr>
<td>25</td>
<td>Khmer (ខ្មែរ)</td>
<td>Abugida</td>
<td>11.4</td>
</tr>
<tr>
<td>26</td>
<td>Greek (ΙΟΥΠΙΤΕΡ)</td>
<td>Alphabet</td>
<td>11</td>
</tr>
<tr>
<td>27</td>
<td>Lontara</td>
<td>Abugida</td>
<td>7.6</td>
</tr>
<tr>
<td>28</td>
<td>Tibetan (བོད)</td>
<td>Abugida</td>
<td>5</td>
</tr>
<tr>
<td>29</td>
<td>Georgian (ქართული)</td>
<td>Alphabet</td>
<td>4.5</td>
</tr>
<tr>
<td>30</td>
<td>Modern Yi (毅)</td>
<td>Syllabary</td>
<td>4</td>
</tr>
<tr>
<td>31</td>
<td>Mongolian (монгол)</td>
<td>Alphabet</td>
<td>2</td>
</tr>
<tr>
<td>32</td>
<td>Tifinagh</td>
<td>Abjad</td>
<td>1</td>
</tr>
<tr>
<td>33</td>
<td>Syriac</td>
<td>Abjad</td>
<td>0.4</td>
</tr>
<tr>
<td>34</td>
<td>Thaana</td>
<td>Abugida</td>
<td>0.35</td>
</tr>
<tr>
<td>35</td>
<td>Inuktitut (ᐃᐃᖃᓗᐃᔭᖃᑎᐸ)</td>
<td>Abugida</td>
<td>0.035</td>
</tr>
<tr>
<td>36</td>
<td>Cherokee (ᏣᎳᎦ)</td>
<td>Syllabary</td>
<td>0.02</td>
</tr>
</tbody>
</table>
Type of the dominant writing system

A global view on world's writing systems
Writing direction
Final remarks
• diversity, 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-syllabaries
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 (phoneme, word) but subscribes to the principle of the primacy of speech)