The evolution of writing systems: Empirical and cross-linguistic approaches workshop (AG5) @ DGfS2020, Universität Hamburg, Germany; 4-6 March, 2020

‘Evolution’ of writing systems in terms of typological and other criteria: Cross-linguistic observations from the German and Japanese writing systems

Terry Joyce
Tama University, Japan
terry@tama.ac.jp

Dimitrios Meletis
University of Graz, Austria
dimitrios.meletis@outlook.com
Overview

- Opening remarks
- Selective sample of writing system (WS) typologies
- Alternative criteria for evaluating WSs
- Observations from German (GWS) + Japanese (JWS)
- Closing remarks
Erring towards understatement, Gnanadesikan (2017: 15) notes, "there is, in general, significant variation in the basic terminology used in the study of writing systems."

Indeed, as Meletis (2018: 73) observes regarding the differences between the concepts of WS and orthography, "these terms are often shockingly misused as synonyms, or writing system is not used at all and orthography is employed instead."

Similarly, Joyce and Masuda (in press) seek to differentiate between the elusive trinity of terms at heart of WS research; namely, script, WS, and orthography, with particular reference to the JWS."
Opening remarks 2: Our working definitions

\( \text{WS}_1 \) [Schrifttyp]: Abstract relations (i.e., morphographic, syllabographic, + phonemic), as focus of typologies.

\( \text{WS}_2 \) [Schriftsystem]: Common usage for signs + conventions of given language, such as GWS + JWS.

\text{Script} [Schrift]: Set of material signs for specific language.

\text{Orthography} [Orthographie]: Mediation between script + WS, but often with prescriptive connotations of \textbf{correct writing}.

\textbf{Graphematic representation}: Emerging from grapholinguistic approach, a neutral (ego preferable) alternative to orthography.

\textbf{GWS}: Use of extended alphabetic set, as used to represent written German language.

\textbf{JWS}: Collection of component scripts (kanji, hiragana, katakana, and rōmaji), as used to represent written Japanese language (Joyce & Masuda 2018).
Opening remarks 3: Reexamining WS typologies

WS typology aims to elucidate key characteristic(s) of WSs, but given their complex nature, general enterprise is not without inherent quagmires (Joyce & Borgwaldt 2011; Joyce 2016).

- Early fallacious assumptions of teleological transitions (Gelb 1952, Daniels 1990).
- Mainly synchronic in nature, the characteristic of focus is invariably the dominate level of graphematic representation. However, such levels and their associated spelling principles (representational mappings) combine in complex ways. In reality, most WSs are - to varying degrees - mixed in nature!

By briefly reflecting on some alternative criteria, this talk aspires to contribute to the ongoing reexamination of WS typology as a crucial tool for understanding WSs, from both synchronic and diachronic perspectives.
Overview

- Opening remarks
- Selective sample of writing system (WS) typologies
- Alternative criteria for evaluating WSs
- Observations from German (GWS) + Japanese (JWS)
- Closing remarks
No Writing: *Pictures*

Forerunners of Writing: *Semasiography*

1. Descriptive-Representational Device
2. Identifying-Mnemonic Device

Full Writing: *Phonography*

   - Elamite
   - Hurrian etc.
   - West Semitic (Phoenician)
   - (Hebrew)
   - (Aramaic)
   - etc.
   - Cypriote
   - Phaistos?
   - Byblos?

2. *Syllabic*: Greek
   - Aramaic (vocalized)
   - Hebrew (vocalized)
   - Latin
   - Indic etc.

WS typologies 1: Gelb (1952, 1963)

writing

- semasiographic
- glottographic

  - logographic (based on polymorphemic unit (e.g. word))
  - phonographic
    - syllabic
    - segmental
    - featural

- morphemic
Undoubtedly, Daniels’ (1990, 2001, 2009, 2018) classification has been one of the most influential of the last two decades. It initially consisted of 6 categories, but 2018 version drops (6)!

(1) Logosyllabary (morphosyllabary)
(2) Syllabary,
(3) Abjad (Semitic-type script), where each character stands for a consonant
(4) Alphabet (Greek-type script)
(5) Abugida (Sanskrit-type script), where each character stands for a consonant accompanied by a particular vowel, with other vowels indicated by additions to the consonant symbol
(6) Featural, where shapes of the characters correlate with phonetic features of designated segments.
### WS typologies 4: Gnanadesikan (2017)

#### Proposal for phonemic WSs (Table 1: 28)

<table>
<thead>
<tr>
<th>Category</th>
<th>Values</th>
<th>Term</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characters (basically) represent segments</td>
<td>Yes</td>
<td>Segmentary/Phonemic script/Segmental script</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>Other (e.g., Syllabary)</td>
</tr>
<tr>
<td>Other structures represented (other than those in ‘higher-order structures’ below)</td>
<td>Features</td>
<td>Featural</td>
</tr>
<tr>
<td></td>
<td>Moras</td>
<td>Moraic</td>
</tr>
<tr>
<td></td>
<td>None</td>
<td>(omit)</td>
</tr>
<tr>
<td>Higher-order structure represented</td>
<td>Peak/margin</td>
<td>Ākšarik</td>
</tr>
<tr>
<td></td>
<td>Syllables</td>
<td>Syllabically arranged/spaced</td>
</tr>
<tr>
<td></td>
<td>None</td>
<td>Linear</td>
</tr>
<tr>
<td>Inclusion of vowels</td>
<td>All</td>
<td>Fully vowelled</td>
</tr>
<tr>
<td></td>
<td>Most</td>
<td>Mostly vowelled</td>
</tr>
<tr>
<td></td>
<td>Some</td>
<td>Partially vowelled</td>
</tr>
<tr>
<td></td>
<td>None</td>
<td>Consonantal</td>
</tr>
</tbody>
</table>
Overview

- Opening remarks
- Selective sample of writing system (WS) typologies
- **Alternative criteria for evaluating WSs**
- Observations from German (GWS) + Japanese WS (JWS)
- Closing remarks
WS typologies are descriptive in nature, but conventional category labels are also employed for (reductive) evaluations!

Jones & Mooney (2017: 13): Overall, it is argued that morphographic systems are inferior to phonographic ones.

Beyond representational mapping, alternative criteria are commonly used for evaluations. However, they could also be beneficially employed for typological purposes.

- Indeed, any criterion that affords a classification of WSs into contrastive categories (such as spaced vs. unspaced) can serve as tenable basis for a typology.
### Alternative criteria 2: Criteria for WS evaluation 1

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>motivation for the learner</td>
<td>maximum distinctiveness</td>
<td>linguistically sound</td>
</tr>
<tr>
<td>representation of speech</td>
<td>size of the graph(emic)ic inventory</td>
<td>acceptable to all stakeholders</td>
</tr>
<tr>
<td>ease of learning</td>
<td>cognitive salience</td>
<td>usable</td>
</tr>
<tr>
<td>transfer</td>
<td>maximum naturalness</td>
<td></td>
</tr>
<tr>
<td>ease of reproduction</td>
<td>inner consistency</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>convenience</td>
<td>linguistic distance</td>
</tr>
<tr>
<td>tools</td>
<td>spatial arrangement and non-linearity</td>
</tr>
<tr>
<td>general applicability + linguistic fit</td>
<td>visual uniformity + complexity</td>
</tr>
<tr>
<td>expressive power</td>
<td>historical change</td>
</tr>
<tr>
<td>simplicity</td>
<td>spelling constancy despite morphophonemic alternation</td>
</tr>
<tr>
<td>stability through time</td>
<td>omission of phonological elements</td>
</tr>
<tr>
<td>monochrome coding</td>
<td>allography</td>
</tr>
<tr>
<td></td>
<td>dual purpose letters</td>
</tr>
<tr>
<td></td>
<td>ligaturing</td>
</tr>
<tr>
<td></td>
<td>inventory size</td>
</tr>
</tbody>
</table>
## Alternative criteria 3: Criteria for WS evaluation 2

<table>
<thead>
<tr>
<th><strong>Venezky (1977)</strong></th>
<th><strong>Bauernschmidt (1980)</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>mechanically suited for the language it is to reflect</td>
<td>linguistic factors</td>
</tr>
<tr>
<td>compatible with [...] its social-cultural setting</td>
<td>psycholinguistic factors</td>
</tr>
<tr>
<td>psychologically/pedagogically appropriate for its speakers</td>
<td>&quot;magic of written language&quot;</td>
</tr>
<tr>
<td></td>
<td>native speaker reactions</td>
</tr>
<tr>
<td></td>
<td>optimal inventory of symbols</td>
</tr>
<tr>
<td></td>
<td>overuse of symbols</td>
</tr>
<tr>
<td><strong>Rogers (1995)</strong></td>
<td>sociolinguistic factors</td>
</tr>
<tr>
<td>linguistic</td>
<td>symbol value</td>
</tr>
<tr>
<td>psychological</td>
<td>adjustments for dialects</td>
</tr>
<tr>
<td>cultural</td>
<td>unity of language families</td>
</tr>
<tr>
<td>technical</td>
<td>prestige, numbers, and so forth</td>
</tr>
<tr>
<td></td>
<td>established alphabets</td>
</tr>
<tr>
<td></td>
<td>government agencies</td>
</tr>
<tr>
<td></td>
<td>transfer value</td>
</tr>
<tr>
<td></td>
<td>practical factors</td>
</tr>
</tbody>
</table>
Alternative criteria 4: Systematizing for WS evaluation

[Meletis 2019: Chapter 3]

- **processing fit**
  - UNIVERSAL
  - how suitable WS is for physiological and cognitive processing

- **linguistic fit**
  - TYPOLOGICAL
  - how well WS maps to other levels (phonological, morphological) of given language (type)

- **sociocultural fit**
  - SYSTEM-SPECIFIC
  - how given WS meets needs of given literate culture/community

Crucial (cyclical) interaction for **diachronic development**: processing fit affects WS structure (linguistic fit), which, in turn, affects processing fit, …
Overview

- Opening remarks
- Selective sample of writing system (WS) typologies
- Alternative criteria for evaluating WSs
- Observations from German (GWS) + Japanese (JWS)
- Closing remarks
Observation 1: GWS 1: Representational mapping

Sampson (2018) notes WSs tend to evolve from phonetically-based when ‘young’ → **lexically-distinctive** as they ‘mature’. Written forms are less dependent on spoken forms, naturally, with huge implications for reading (and writing) processes.

Thus, even though segmental in nature, syllabic structures are rendered visible in GWS by graphe(ma)tic feature of **length** (most notably represented by ascenders and descenders).

Diachronically, visually long shapes have been ousted from the syllable nucleus position (Fuhrhop & Schmidt 2014) in order to render syllable boundaries more transparent.

- long |jj| → non-long |i|, as in <jhn> → <ihn>, when syllable nucleus
- Inversely, <iung> → <jung>, with the long |jj| representing the syllable boundary (Fuhrhop & Schmidt 2014: 559).
Observation 2: JWS 1: Mixed WS

漢字 /kan-ji/ (lit. ‘Chinese characters’): Morphographic
Represent both native-Japanese and Sino-Japanese content words, including nouns, stems of verbs and of some adjectives, and some adverbs.

ひらがな /hira-ga-na/: Syllabographic
Represent functional words, inflectional elements of verbs and some adjectives, grammatical case marks and conjunctions.

カタカナ /kata-ka-na/: Syllabographic
Represent foreign-Japanese (but not Chinese words), foreign names, animal and plant species names, onomatopoeia, emphasis and as glosses.

Rōmaji: Phonemic
Represent foreign words and names, particularly in advertising and mass media.
Observation 3: GWS 2: Word spacing 1

Word spacing supports, at the graphetic (visual) level, *saccade programming*, and, at the graphematic (linguistic) level, *lexical segmentation*.

- Introduction of word spacing for alphabets heralded the breakthrough to silent reading – it was a response to processing needs.
- Experimental evidence suggests that deletion of word spacing (in absence of alternative boundary cues) is detrimental to reading processes for WSs that are standardly spaced; in contrast, the introduction of word spacing is not necessarily beneficial, where unspaced is the standard (e.g., Chinese, Thai).
- Word spacing is not the only way to provide boundary cues.
Observation 4: JWS 2: Word spacing 2: Not-spaced!

Joyce & Masuda (2018: 183) has following example of authentic Japanese sentence (Wikipedia entry for encoding standard) with all component scripts (attested, but admittedly rare).

JIS X 0208 (ジス X 0208) は、日本語表記、地名、人名などで用いられる6,879図形文字を含む、主として情報交換用の2バイト符号化文字集合を規定する日本工業規格である。
Observation 4: JWS 2: Word spacing 2: Not-spaced!

Joyce & Masuda (2018: 183) has following example of authentic Japanese sentence (Wikipedia entry for encoding standard) with all component scripts (attested, but admittedly rare).

JIS X 0208 (ジス X 0208) は、日本語表記、地名、人名などで用いられる6,879図形文字を含む、主として情報交換用の2バイト符号化文字集合を規定する日本工業規格である。

Color coding: kanji, hiragana, katakana, rōmaji + numbers


JIS X 0208 is a Japanese Industrial Standard that stipulates a 2-byte encoded character set that is mainly used in information exchange that includes 6,879 graphic characters that are used for Japanese language writing, place names and personal names, etc.
Structural characteristics of WSs are predictive of which aspects are likely to be subjected to standardization and orthographic regulation (Coulmas, 2016).

- 1996 reforms of German orthography included grapheme-phoneme correspondences (Stengel $\rightarrow$ Stängel, Kuß $\rightarrow$ Kuss), word divisions (radfahren $\rightarrow$ Rad fahren), capitalization (in bezug auf $\rightarrow$ in Bezug auf), and punctuation (no comma before und when followed by the main clause).
- 2006 reforms/amendments revoked some changes, while others made optional (both old + new variants permissible).
- However, optionality is not desirable to most users (Nerius 2007; Meletis 2019); significant role of orthography in prescribing what to regard as correct from among possible alternative graphematic representations.
Observation 6: JWS 3: Orthographic regulation 2

**Kanji:** From mid-20-century, Japanese government issuing guidelines aimed at limitations (plus some simplification)

<table>
<thead>
<tr>
<th>'List of kanji for general use’ guidelines</th>
<th>Date</th>
<th>Kanji</th>
</tr>
</thead>
<tbody>
<tr>
<td>当用漢字表 /tō-yō-kan-ji-hyō/</td>
<td>1946</td>
<td>1,850</td>
</tr>
<tr>
<td>常用漢字表 /jō-yō-kan-ji-hyō/</td>
<td>1981 Oct</td>
<td>1,945</td>
</tr>
<tr>
<td>Jōyō kanji list revision [-5 +196 = 191]</td>
<td>2010 Nov</td>
<td>2,136</td>
</tr>
<tr>
<td>教育漢字 /kyō-iku-kan-ji/ ‘education kanji’ for Grades 1-6</td>
<td></td>
<td>(1,006)</td>
</tr>
<tr>
<td>Remainder taught during high school</td>
<td></td>
<td>(1,130)</td>
</tr>
</tbody>
</table>

**Kana:** Generally consistent, but modern kana usage influenced by some historical conventions;

- 3 kana inconsistent as case particles: は /ha/ → /wa/ (topic);
  へ /he/ → /e/ (destination);
  を /wo/ → /o/ (object; sole usage).
- Morphology influences: 鼻血 /hana-ji/ ‘nose bleed’ - はなぢ /hana-di/ (not はなじ), as 血 ‘blood’ usually as ち /chi/.
Overview

- Opening remarks
- Selective sample of writing system (WS) typologies
- Alternative criteria for evaluating WSs
- Observations from German (GWS) + Japanese WS (JWS)
- Closing remarks
Closing remarks 1: Alternative criteria

With laudable aim of elucidating key characteristics, WS typology have generally focused on representational mapping. However, it is vital to keep in mind that notions of morphographic, syllabographic + phonemic are idealizations.

Given that most WSs are mixed in nature, WS typological accounts should develop ways of more adequately capturing their mixed and complex nature. Against that background, this talk advocates the exploration of alternative criteria – either in own right or in combination – which can potentially further enhance descriptions of WSs in all their glory. Examples singled out for illustration relate to Meletis’ (2019) framework for 3 kinds: linguistic fit, processing fit and sociocultural fit.
Closing remarks 2: Linguistic fit

Again, linguistic fit is assessed for individual WS, but usually employs WS typology notion of representational mapping.

JWS generally acknowledged as most problematic case (Joyce 2011, 2016; Sproat 2010), due its component scripts. Orthographic convention is known as 漢字かな混じり文 /kan-ji.kana.ma.jiri.bun/ (lit. kanji + kana mixed writing). Mix is also a factor for ubiquity of graphematic variation in JWS (Joyce, Hodošček, & Nishina 2012; Joyce & Masuda in press).

Script sensibilities (Joyce & Masuda in press)
私, わたし, ワタシ, watashi /watashi/
‘I’ [both genders, but more commonly female]
Closing remarks 3: Processing fit

Also singled out presence/absence of word spacing, noting its presence aids saccade programming and lexical segmentation within phonemic WSs. However, even such as ‘minimal’ shift can have potentially far-reaching implications, for the development of punctuation and hyphenation rules.

**JIS X 0208 (ジス X 0208) is a Japanese industrial specification that defines the set of 6,879 graphic characters used in Japanese, including mainly those for information exchange.**

In non-spacing WS such as JWS, line-wrapping frequently occurs within word.

用いられる /mochi.irareru/ ‘that are used’ (plain passive)

2バイト符号化文字集合 ‘2-byte encoded character set’
Orthographic reforms are generally implemented on the grounds that they will help improve general literacy levels. However, on the complex interactions between linguistic, social, political, and economic aspects, Coulmas (2013: 124) remarks that the perpetuation of inconsistencies and the introduction of new ones seem an inevitable side-effect of writing reforms.

Also in connection with script engineering, Sproat’s (2010: 136) correlation analyses indicate how one factor that does not correlate with literacy is the complexity of the script.
Given the complexities of natural WSs, it seems rather foolhardy to continuing overly focusing on single dimensions, such as representational mapping, no matter how well founded it might be from core linguistic perspectives. There would seem to be considerable merit in considering alternative criteria – either independently or in tandem – in order to better understanding WSs, both from synchronic and diachronic perspectives.

Danke für Ihre Aufmerksamkeit
ご清聴ありがとうございます
Thank you for your kind attention


References 2


