

A proposal for a formalized, expandable approach to the taxonomy of writing systems

Sven Osterkamp & Gordian Schreiber
(Ruhr University Bochum)

AWLL13 | 21–23 October 2021

Outline

- Issues in current taxonomies
- A formalized, expandable approach

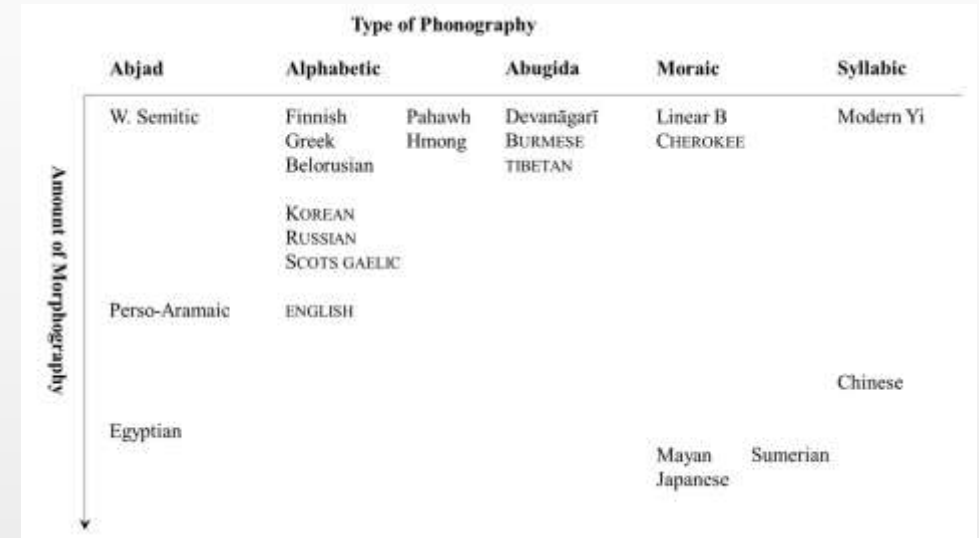
Issues in current taxonomies

‘Messiness’

- ‘messy’ / ‘mixed’ / ‘hybrid’ nature of writing systems widely acknowledged, but notion as such not uniform (e.g. Sampson 1985, Coulmas 2003, Rogers 2005, Daniels 2018)
 - co-existence of phonography and morphography
 - reference to different linguistic levels, e.g. when the choice of phonograms requires morpheme-specific knowledge (cf. Osterkamp & Schreiber 2021a)
 - anything going beyond 1:1 correspondences (cf. Osterkamp & Schreiber 2021b)

Planar taxonomies

- different kinds of phonography treated as mutually exclusive (Joyce 2016: 296)
- cases in between categories limited here to ‘adjacent’ ones
- cf. already Daniels (1990): *augmented abjads* and *alphabets* (Ugaritic, Coptic)



Rogers' (2005) classification of writing systems,
based on Sproat (2000)

Categories

- categories commonly posited on the basis of only a fraction of the known writing systems
- often fail to accommodate cases situated in between
- some categories conflate features belonging to different and in principle independent levels: *what* is represented as opposed to *how*?
 - e.g. alphasyllabary: “vowels are denoted by subsidiary symbols not all of which occur in a linear order that is congruent with the temporal order of speech” (Daniels & Bright 1996: xxxix)

Labels

- categories are not labeled in a uniform fashion (Joyce 2016: 295)
 - “linguistic levels”: logosyllabary, syllabary
 - “exemplar names”: abjad, abugida, alphabet
 - “grapheme structure”: featural
- sometimes a descriptive account is provided, but no classification or label

A formalized, expandable approach

A formalized, expandable approach

- focusing on the classification of typologically homogeneous subsystems coexisting in writing systems (Osterkamp & Schreiber 2021a)
- side-effect: increasing the visibility of non-dominant subsystems
- basic distinction into segmental, syllabic and morphemic systems, complemented by an open number of subtypes, building and expanding on Gnanadesikan (2017) and Poser (1992, 2004) for segmental and syllabic systems respectively

Overview of symbols used

form: boundedness of graphs

- free / bound graphs

X/x

function: mapping on linguistic units

- phono- / morphograms
- consonants, vowels, ...
- long / short; specific, inherent vowels
- unwritten elements
- suppressed portions
- free / bound morphemes

P/M, p/m

C, V, CV, VC, CVC, ...

$\bar{V}/\check{V}$, $\bar{v}/\check{v}$; V_x , V_i (also C_i)

[C], [V], ...

$C(V_i)$, $(C_i)V$, ...

F/B, f/b

Segmental systems: options for writing /CV/

English

alphabet

<a.l.**ph**.a.b.e.t>

/æ**f**əbɛt/

<C.V>

e.g. <C> = written unit mapped onto /C/, one or more free graphs

Segmental systems: options for writing /CV/

Arabic
(unvocalized)

نام

<n.ā.m>

/naːma/ ‘to sleep’

<C.Ṫ>, <C[Ṫ]>

Arabic
(vocalized)

نَامَ

<n^(a).ā.m^a>

/naːma/ ‘ditto’

<C(ṽ).Ṫ>, <Cṽ>

Segmental systems: options for writing /CV/

Devanāgarī
(Sanskrit)

वन

<va.na>

/vāna/ ‘forest’

<CV_i>

or: <C[V_x]>

वेद

<v(a)^e.da>

/veda/ ‘knowledge’

<C(V_i)v>

or: <Cv>

Segmental systems: options for writing /CV/

'Phags-pa
(Old Mandarin)

𐞪

<**na**>

/**na**/ 捺 'press down'

<CV_i>

or: <C[V_x]>

𐞫

<**n(a).o**>

/**no**/ 那 'that'

<C(V_i).V>

or: <C.V>

Syllabic systems: options for writing /CVC/

Old Japanese

之良末世婆

<si.ra.ma.se.ba>

<CV>

/siramaseba/ ‘had I known’

Mandarin
Chinese

华盛顿

<huá.shèng.dùn>

<CVC>

/huáshèngdùn/ ‘Washington’

Syllabic systems: options for writing /CVC/

Akkadian
cuneiform



<**ka**.**lu**>

/**kal**lu/ ‘bowl’

<CVC>



<**ka**.**(a)****lu**>

/**kal**lu/ ‘bowl’

<CV_x.(V_x)C>
 (“demi-syllabic”)

Syllabic systems: options for writing /CVC/

Hiragana
(Mod. Japanese)

よん
<**yo.N**>
/y**oN**/ ‘four’

<CV.C>
 (“moraic”)

early Hiragana
(Classical Jp.)

よむたる
<**yo.m(u).ta.ru**>
/y**om**daru/ ‘has recited’

<CV.C(V_x)>
 (“proto-moraic”)

もたり
<**mo.ta.ri**>
/m**ottari**/ ‘have, own’

<CV[C]>
 (“head moraic”)

Syllabic systems: options for writing /CVC/

Devanāgarī
(Sanskrit)

इदम्

$\langle i.da.m(a)^{*v} \rangle$
/idam/ 'this'

$\langle CV_i.C(V_i)^{*v} \rangle$
or: $\langle C[V_x].C^*v \rangle$

किम्

$\langle k(a)^i.m(a)^{*v} \rangle$
/kim/ 'what'

$\langle C(V_i)v.C(V_i)^{*v} \rangle$
or: $\langle Cv.C^*v \rangle$

Syllabic systems: options for writing /CVC/

Lepcha

ᱫᱤᱨ

<ka.**k**(a)^u>

/kaku/ 'eight'

<C(V_i)v>

or: <Cv>

ᱫᱤᱞ

<**k**(a)^{u-p}>

/kup/ 'child, small'

<C(V_i)v-c>

or: <Cv-c>

Syllabic systems: options for writing /CVC/

Bopomofo
(Mandarin)

ㄏㄜˊ

<**h.e**²>

/h^é/ 和 ‘and’

<C.V>

ㄏㄣˇ

<**h.en**³>

/h^ěn/ 很 ‘very’

<C.VC>

(“onset/rhyme-based”)

Morphemic systems: options for writing words

Mandarin
Chinese

看

<**KÀN**>

/**kàn**/ ‘to look at’

<M>

Mod. Japanese

見る

<MI.**ru**>

/mi**ru**/ ‘to look at’

<M.P>

Morphemic systems: options for writing words

(W)okototen
(Classical Jp.)



[入(り)たまふ]

<IRI^{TAMAFU}>

<Mm>

/irit**amafu**/ ‘deigns to enter’



[昔し]

<MUKA(SI)^{si}>

<Mp>

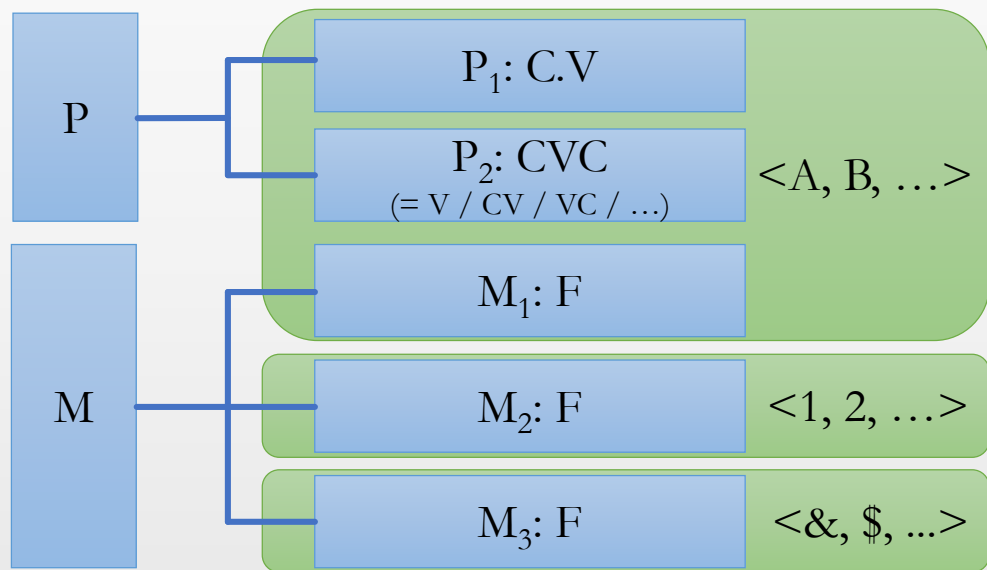
/muk**asi**/ ‘in the old days’

Script – subsystem

- script: inventory of graphs considered *graphically* homogeneous
- subsystem: (subset of a) script that is in a *functionally* homogenous fashion
- subsystems often overlap, but may also be in complementary distribution:
 - unvocalized Arabic: $C.\bar{V} / C[\check{V}]$
 - Devanāgarī: $CV_i / C(V_i)_v$ or $C[V_x] / C_v$

English

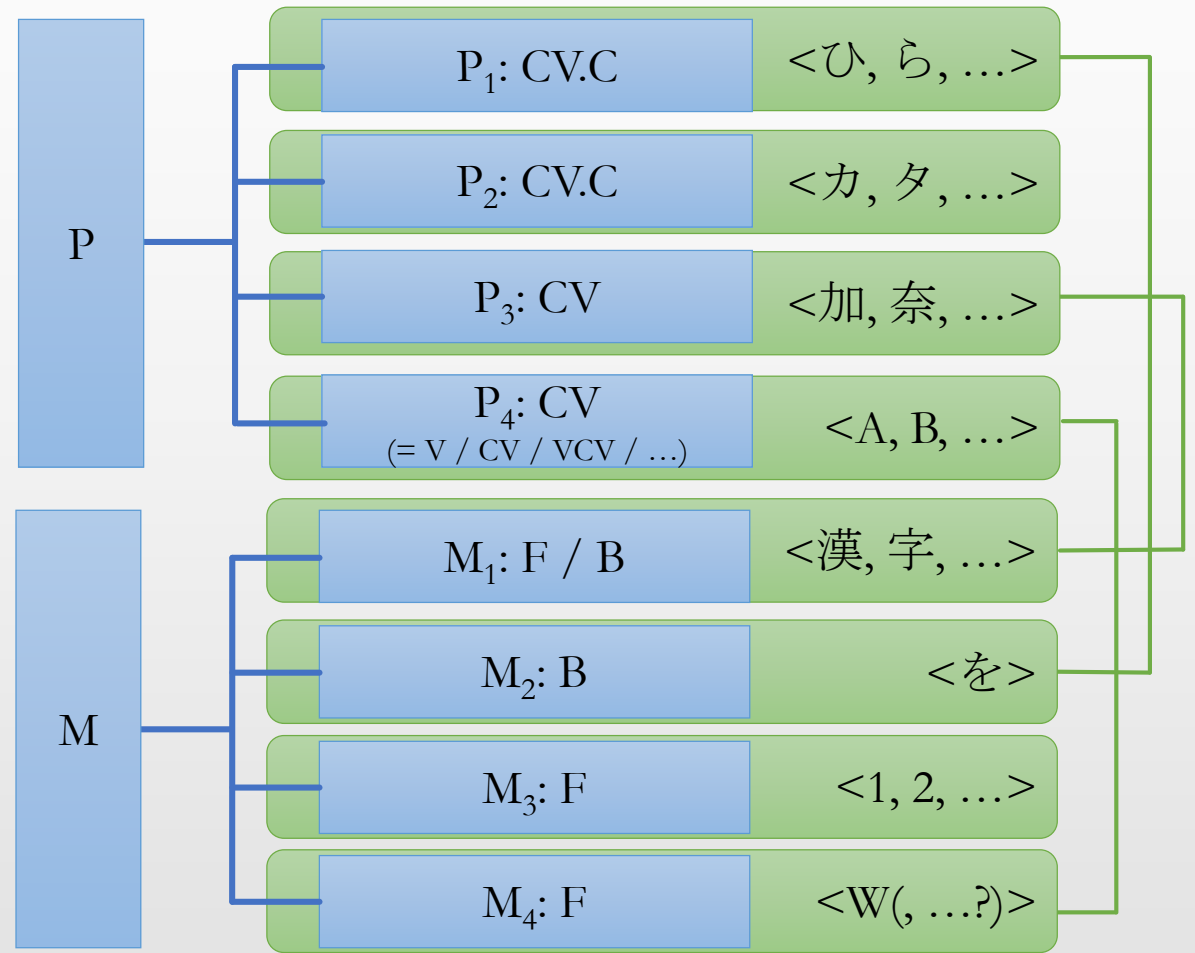
P_1 (dominant) / P_2 / $M_{1/2/3}$ / $M_2.P_1$



P_2 : <pdf, IQ, ...>, M_1 : <e.g., i.e., ...>
(cf. Gelb 1963: 16)

Modern Standard Japanese

$P_{1/2/3/4}(.M_2)$ / $M_{1/3}(.P_1)(.M_2)$ / ...



...

P_3 : <CD> (*shīdi*), M_4 : <W> (*daburu* 'double')

Outlook: Further parameters

- graphical configuration
 - directionality of writing
 - non-sequential writing/reading order (cf. Schreiber 2022)
 - linking and separation (syllable/word/... division)
- determinatives, tonal notation
- mapping of graphic manipulations (rotation, reflection, resizing) on linguistic units? or better treated as an issue of etymography?

References

- Coulmas, Florian (2003). *Writing Systems. An Introduction to their Linguistic Analysis*. Cambridge: Cambridge UP.
- Daniels, Peter T. & William Bright (eds.) (1996). *The World's Writing Systems*. New York: Oxford UP.
- Gelb, I. J. (1963). *A Study of Writing*. Revised Edition. Chicago: Chicago UP.
- Daniels, Peter T. (1990). Fundamentals of Grammatology. *Journal of the American Oriental Society* 110.4: 727–731.
- Daniels, Peter T. (2018). *An Exploration of Writing*. Sheffield: Equinox.
- Gnanadesikan, Amalia E. (2017). Towards a Typology of Phonemic Scripts. *Writing Systems Research* 9.1: 14–35.
- Osterkamp, Sven & Schreiber, Gordian (2021a). Challenging the Dichotomy Between Phonography and Morphography: Transitions and Gray Areas. In: Yannis Haralambous (ed.): *Grapholinguistics in the 21st Century [...] 2020, Proceedings, Part I*. Fluxus Editions, pp. 47–82.
- Osterkamp, Sven & Schreiber, Gordian (2021b). <Th>e Ubi<qu>ity of Polygra<ph>y and its Significan<ce> for <th>e Typology of <Wr>iti<ng> Systems. *Written Language & Literacy* 24: 171–197.
- Poser, William J. (1992). The Structural Typology of Phonological Writing. Handout (Linguistic Society of America, Philadelphia, 12 Jan 1992).
- Poser, William J. (2004). Phonological Writing and Phonological Representation. Handout (Harvard University, 10 Dec 2004).
- Rogers, Henry (2005). *Writing Systems. A Linguistic Approach*. Malden: Blackwell.
- Sampson, Geoffrey (1985). *Writing Systems. A Linguistic Introduction*. London: Hutchinson.
- Schreiber, Gordian (2022). *Japanese Morphography. Deconstructing hentai kanbun*. Leiden: Brill.
- Sproat, Richard (2000). *A Computational Theory of Writing Systems*. Cambridge: Cambridge UP.