

Syllables and Syllabaries: What Writing Systems Tell Us about Syllable Structure

Becoming literate requires becoming consciously aware of some subset of the linguistic structures in one's language. The typology of writing systems can therefore inform us about the range of structures that can be raised to conscious awareness, which are by implication a subset of the structures that can be considered psychologically "real" and active in a given language. Most early writing systems, as well as later inventions by preliterate individuals, are syllabaries (or logosyllabaries), not alphabets. Presumably then, the syllable is a linguistic structure that is available to both explicit and implicit knowledge; indeed, explicit awareness of syllables has been shown to developmentally precede explicit awareness of phonemes. However, the sign inventory of most syllabaries is a poor match for the syllabic inventory of the relevant languages, apparently belying a relationship between the syllabary and the language's phonology. Do syllabaries actually encode syllables, then, or merely arbitrary groupings of segments?

I argue that syllabaries do in fact provide evidence of syllabic units and of many properties of syllables, such as onset-over-coda preference, extrasyllabicity of /s/, moras, and the onset/rime distinction. First, many syllabaries do encode syllables, though they may distort the linear order of segments. Akkadian cuneiform includes symbols for units of CV, V, VC, and CVC. That these are *syllables* rather than segmental strings is evident from the fact that a CVCV string is always written CV-CV, never CVC-V, except at a morpheme boundary (1). Nor are there any CCV signs, a type of syllable which does not occur in Akkadian, but a segmental string that does occur (word internally).

In Akkadian, many $C_1V_2C_3$ syllables are written $S_{12}-S_{23}$; the identity of the two vowels serves to show that there is a single nucleus (and thus syllable). In Mayan (2), $C_1V_2C_3$ is usually written $S_{12}-S_{32}$. A difference in the vowels means that the syllable's nucleus is long, aspirated, or glottalized. Thus in Mayan the vowels in both symbols refer to the same syllable, the second sometimes adding information about it.

In Linear B (3), syllable codas are not notated, but both members of a complex onset are, through repetition of the vowel—a classic case of onset preference over codas. Omission of the (dispreferred) coda consonant serves to keep the structure of the (preferred) onset clear: otherwise $S_{12}-S_{32}$ would be ambiguous between $C_1V_2C_3$ and $C_1C_3V_2$. In the Cypriote syllabary (4), most coda consonants are notated, but codas are followed by a dummy *-e* and so the ambiguity is still avoided. Thus syllabaries steadfastly distinguish onset consonants from coda.

The Cherokee syllabary (5) acknowledges the extrasyllabicity of /s/ with the only phonemic symbol in the script. Linear B (6) omits an extrasyllabic /s/, unlike true onsets.

Some "syllabic" writing systems are actually moraic, lending support for the validity and accessibility of the moraic internal structure of syllables in these languages. In Japanese kana (7), all heavy syllables receive two symbols. Vai (8) is similar.

A few writing systems break syllables into onsets and rimes. In Bopomofo (9), consonantal onsets receive one symbol, a medial glide (if present) another, and the main vowel, diphthong, or vowel-nasal combination another. In the Pahawh Hmong script (10), the rime of a syllable is written first, with tone diacritic, followed by the syllable's onset. Emphasizing the rime, I argue, stems from its tone-bearing nature in these languages.

The world's syllabaries provide evidence for the psychological reality—and accessibility—of various syllabic structures. Other writing systems provide evidence of many other linguistic units, though not, oddly, for the phonological foot.

Example Number	Language/Script	Segment String	Grapheme String
(1)	Akkadian/cuneiform	$C_i V_j - C_k V_l$ $C_i V_j C_k$	$S_{ij} S_{kl} *S_{ijk} S_l$ $S_{ij} S_{jk}$ or S_{ijk}
(2)	Mayan/glyphs	$C_i V_j C_k$ $C_i V_j V_j C_k$ $C_i V_j h C_k$ $C_i V_j ? C_k$	$S_{ij} S_{kj}$ $S_{ij} S_{kx}$ $S_{ij} S_{kx}$ $S_{ij} S_{kx}$
(3)	Greek/Linear B	$C_i V_j C_k$ $C_i C_j V_k C_l$	S_{ij} $S_{ik} S_{jk}$
(4)	Greek/Cypriote	$C_i C_j V_k C_l$	$S_{ik} S_{jk} S_{l-e}$
(5)	Cherokee	$s_i C_j V_k$	$S_i S_{jk}$
(6)	Greek/Linear B	$s_i C_j V_k$	S_{jk}
(7)	Japanese/Kana	$C_i V_j V_j$ $C_i V_j V_k$ $C_i V_j N_l$ $C_i V_j C_x$	$S_{ij} S_j$ $S_{ij} S_k$ $S_{ij} S_l$ $S_{ij} S_x$
(8)	Vai	$C_i V_j V_j$ $C_i V_j V_k$ $C_i V_j \eta_l$	$S_{ij} S_{Gj}$ $S_{ij} S_{Gj}$ $S_{ij} S_l$
(9)	Chinese/Bopomofo	$C_i (G_j) V_k$ $C_i (G_j) V_k V_l$ $C_i (G_j) V_k N_m$	$S_i (S_j) S_k$ $S_i (S_j) S_{kl}$ $S_i (S_j) S_{km}$
(10)	Pahawh Hmong	$C_i (C_j) V_k$ $C_i (C_j) V_k V_l$ $C_i (C_j) V_k \eta_m$	$S_k S_{i(j)}$ $S_{kl} S_{i(j)}$ $S_{km} S_{i(j)}$