

Syllabic Writing and Syllable Structure

Eugene Buckley
University of Pennsylvania

SUNY at Stony Brook — 21 April 2006

Writing Systems and Linguistic Structure

“written language is a product of linguistic awareness” (Aronoff 1985)

- Reflects the categories present in spoken language
 - segment, syllable, word, etc.
- Embodies a theory of language
 - though also affected by cultural history

Traditional Typology of Writing

- Ideographic
 - not truly linguistic; semasiographic
- Logographic
 - most often “morphographic”
- Syllabic
 - or is it moraic?
- Alphabetic
 - segments, potentially just consonants

Moraic or Syllabic?

“In addition to the syllable and the segment, writing systems exist that make use of the constituents *mora* and *rhyme*, and the notion *head*.”

“[Pure] syllabaries are exceedingly rare.”

Poser (1992)

Moraic or Syllabic?

“After all, it is not really the syllable, but rather the light syllable which recurs as a unit of representation in one script type after another.”

Ratcliffe (2001)

Against the Moraic Analysis

- CV systems may include non-moraic elements
- CV systems may fail to write all moraic elements
- CV is not a moraic constituent
- CV is really just a minimal syllable

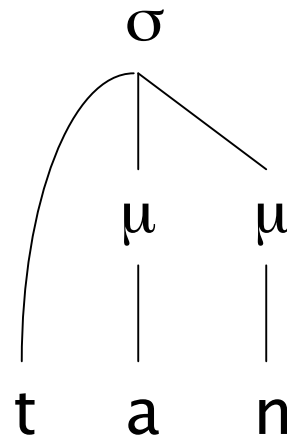
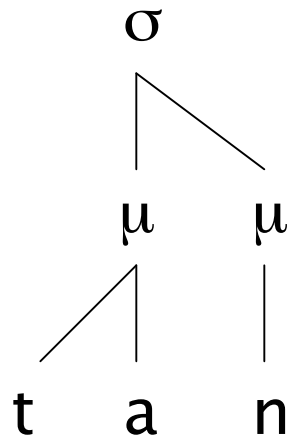
CV as a constituent

The mora-linked onset

Moras in Syllable Structure

Are they
constituents?

Or are they
weight units?



Constituency of CV

If onset links to head mora:

$$[[\text{ta}]_{\mu} [\text{n}]_{\mu}]_{\sigma}$$

If onset links directly to syllable:

$$[\text{t} [\text{a}]_{\mu} [\text{n}]_{\mu}]_{\sigma}$$

Other Constituencies

CV	[[ta] _μ] _σ	[t [a] _μ] _σ
CCV	[[tra] _μ] _σ	[tr [a] _μ] _σ
CV:C	[ta] _μ [an] _μ	t [a] _μ [an] _μ
	[ta] _μ [a] _μ [n] _μ	t [a] _μ [a] _μ [n] _μ

Japanese *kana*

The ultimate moraic example

Japanese *kana* as moraic

“Although the *kana* scripts are often called syllabaries, they are in fact moraic systems. Each symbol in the *kana* scripts represents one mora. Most of these are CV sequences, but final /N/ or /Q/ count as separate morae, and vowel length adds a mora.”

Rogers (2005)

Basic signs in *kana*

Number of signs = number of moras

- (C)V, the core syllables
- N, the “moraic nasal”
- vowel length
 - special sign in *katakana*
 - use second V sign in *hiragana*

Mora matches in *hiragana*

さむらい

<sa-mu-ra-i>

[sa]_μ [mu]_μ [rai]_{μμ}

samurai

“warrior”

おかあさん

<o-ka-a-sa-N>

[o]_μ [kā]_{μμ} [san]_{μμ}

okāsan

“mother”

Mora matches in *katakana*

ネクタイ

<ne-ku-ta-i>

[ne]_μ [ku]_μ [tai]_{μμ}

nekutai

“necktie”

ジーンズ

<zi-L-N-zu>

[jīn]_{μμμ} [zu]_μ

jīnzu

“jeans”

Derived signs in *kana*

Gemination using “Q”

- small version of /tu/ = [tsɯ]
- here, sign = mora

Complex onset in [CyV]

- <Ci> plus small version of <yV>
- here, sign ≠ mora
- similar for some borrowed sequences

Mora matches in *hiragana*

しき

<si-ki>

[š̥i]_μ [ki]_μ

šiki

“command”

しっき

<si-Q-ki>

[š̥ik]_{μμ} [ki]_μ

šikki

“lacquerware”

Mora matches in *katakana*

ロケット

<ro-ke-Q-to>

[ro]_μ [ket]_{μμ} [to]_μ

roketto

“rocket”

ヨーロッパ

<yo-L-ro-Q-pa>

[yō]_{μμ} [rop]_{μμ} [pa]_μ

yōroppa

“Europe”

Mora mismatches in *hiragana*

ちよつと

<ti-yo-Q-to>

[čot]_{μμ} [to]_μ

čotto

“a little”

じゅみょう

<zi-yu-mi-yo-u>

[ju]_μ [myō]_{μμ}

jumyō

“life span”

Mora mismatches in *katakana*

キャッシュ

<ki-ya-Q-si-yu>

[kyaš]_{μμ} [šu]_μ

kyaššu

“cash”

ファミコン

<hu-a-mi-ko-N>

[fa]_μ [mi]_μ [kon]_{μμ}

fami kon

“video game”

Interim Summary 1

- Japanese graphs write CV, which usually corresponds to a mora.
- But parallel techniques are used to write non-moraic differences.
- Other examples:
 - Linear B “onset splitting”
 - Mongolian ‘Phags-pa codas

Ethiopic Alphasyllabary

A systematic syllabary

Ethiopic Alphasyllabary

	ä	u	i	a	e	(i)	o
h	ሀ	ሁ	ሂ	ሃ	ሄ	ህ	ሆ
l	ለ	ሉ	ሊ	ላ	ሌ	ል	ሎ
h	ሐ	ሑ	ሒ	ሓ	ሔ	ሕ	ሖ
m	መ	ሙ	ሚ	ማ	ሜ	ም	ሞ

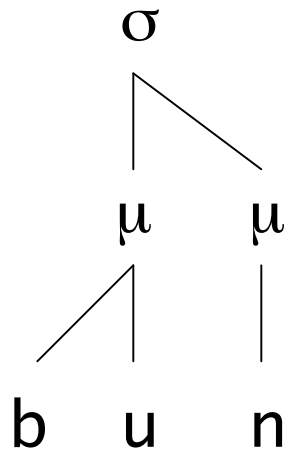
Ethiopic Alphasyllabary

	ä	u	i	a	e	(i)	o
ʔ	አ	ኡ	ኢ	ኣ	ኤ	ኦ	ኰ
b	በ	ቡ	ቢ	ባ	ቤ	ቦ	ቦ
g	ገ	ጉ	ጊ	ጋ	ጌ	ግ	ገ
d	ደ	ዱ	ዲ	ዳ	ዴ	ድ	ደ
	<i>ʔa</i>	<i>bu</i>	<i>gi</i>	<i>da</i>			

Minimal words in Tigrinya

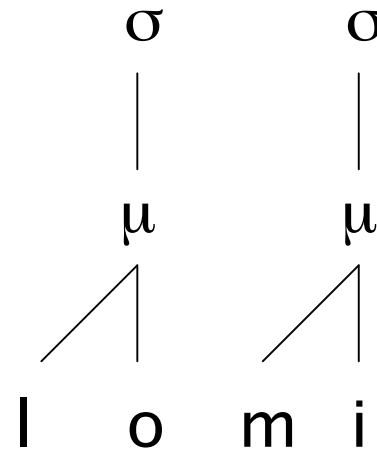
One heavy syllable:

bun 'coffee'



Two light syllables:

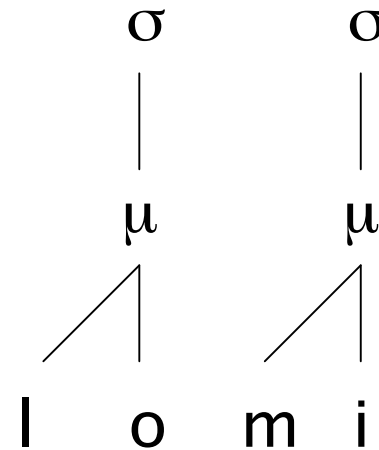
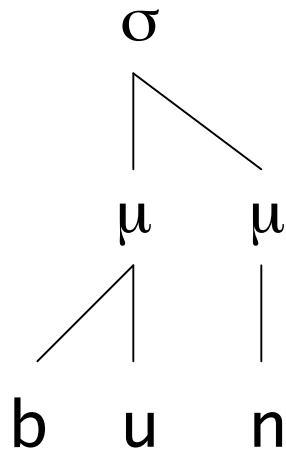
lomi 'today'



Minimal words in Tigrinya

bu-n = ቤን

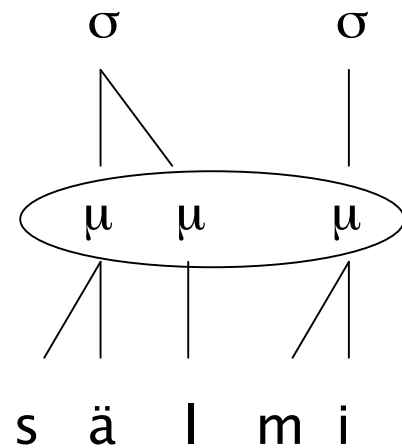
lo-mi = ሎሚ



One sign per mora?

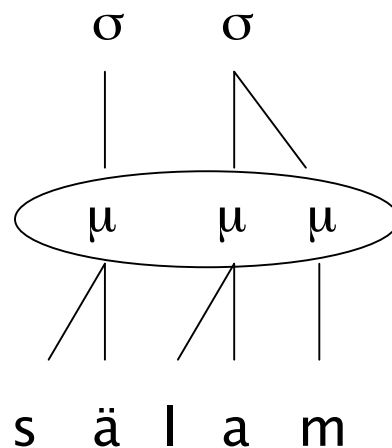
ሰልሚ

sälmi
'muddy soil'



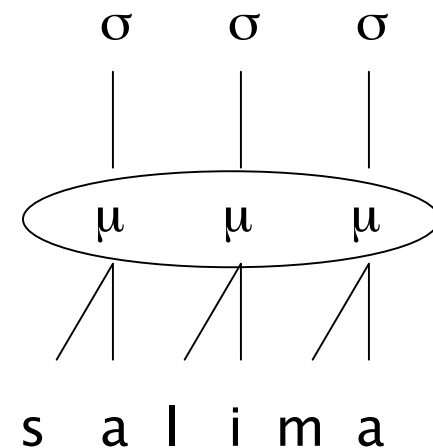
ሰላም

sälam
'peace'



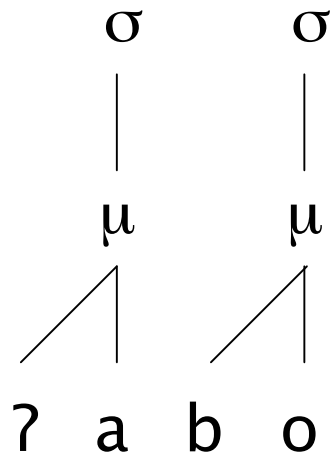
ሳሊማ

salima
'lightning source'

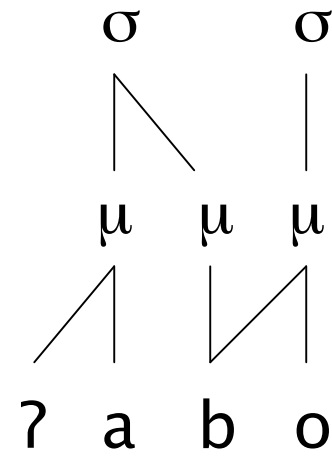


Gemination in Tigrinya

ʔabo 'oleander'



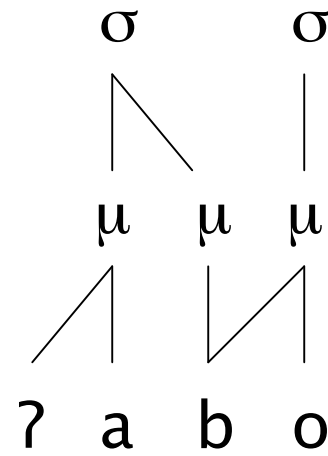
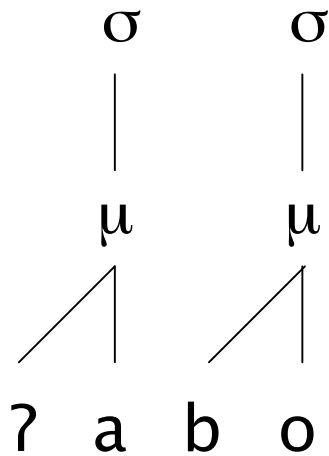
ʔabbo 'father'



Gemination in Tigrinya

ʔabo = ኧቦ

ʔabbo = ኧቦ



Sign ≠ Mora

ኃቦ

ʔa-bo

ʔabo

‘oleander’

ʔabbo

‘father’

not ኃቦቦ ʔa-b-bo

ሰለሙ

sä-lä-mä

sälämä

‘be calm’

sällämä

‘adorn’

not ሰለሰሙ sä-l-lä-mä

Sign ≠ Mora

ይስብሩ

yɨ-sʌ-b-ru

yɨ.sʌb.ru

‘they break’

ይስበሩ

yɨ-sɨ-bʌ-ru

yɨ.sɨb.ʌ.ru

‘they are broken’

Interim Summary 2

- Tigrinya writes CV, a supposed mora, and coda consonants.
- But omits gemination, which is exactly a mora.
- Other examples:
 - Cypriot
 - Akkadian (optional)

Against CV as a mora

Phonological considerations

Moras as Headed Constituents

Both moras and syllables impose minimal sonority threshold on their heads; syllabic \geq moraic.

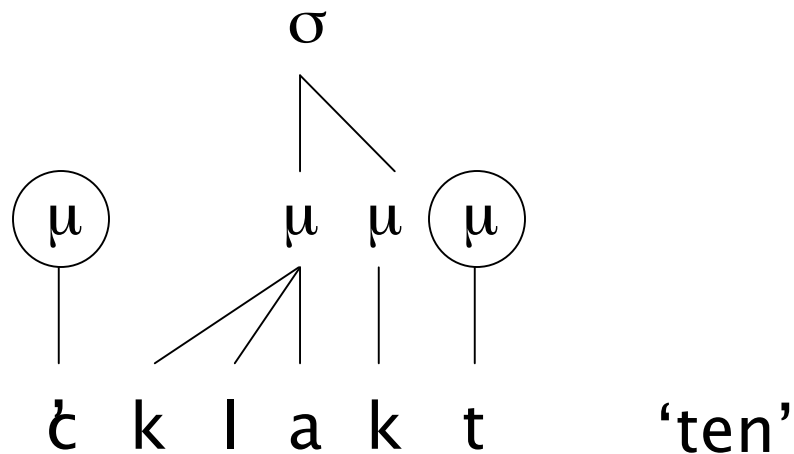
Berber	Serbian	Kwa'kwala
vowel	vowel	σ vowel
liquid	σ liquid	liquid
nasal	nasal	μ nasal
σ, μ obstruent	μ obstruent	obstruent

But ultimately this is orthogonal to onset linking.

Zec (1988, 1995)

Moraic licensing

Bella Coola obstruent clusters often not in syllables, which are rather simple CRV:C.



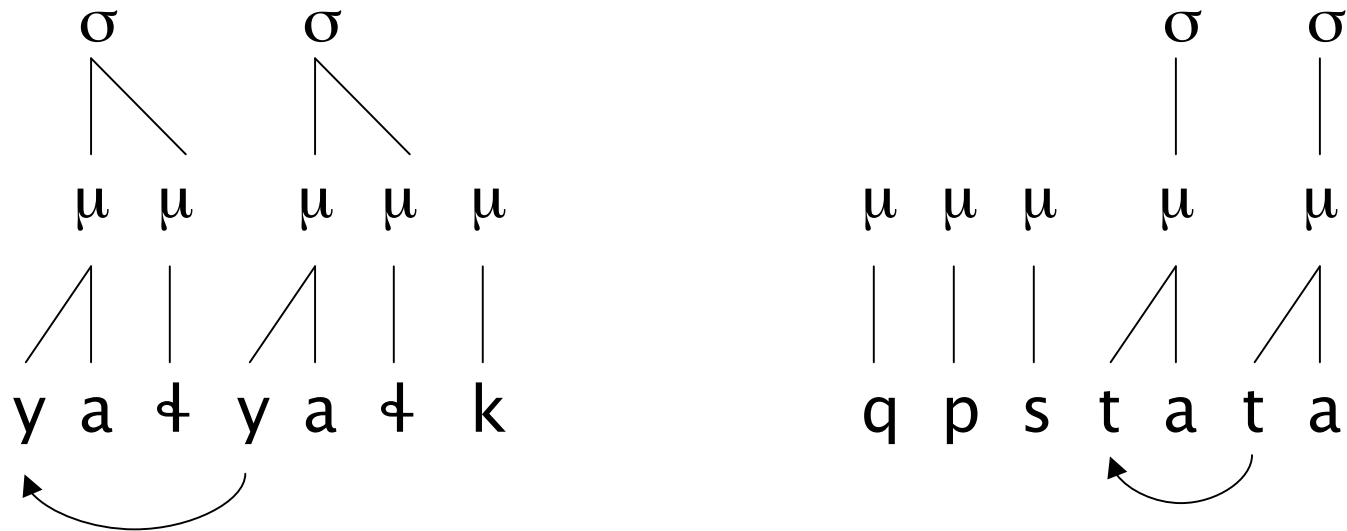
Bagemihl (1991)

Bella Coola reduplication

[mi] [lix ^w]	→	mil-milix ^w -ɬp	‘bear-berry’
[yaɬ] k	→	yaɬ-yaɬk	‘do too much’
s [tan]	→	s-tan-tan-mts	‘mother’
p̥ [ɬa]	→	p̥-ɬa-ɬa	‘wink’
qps [ta]	→	qps-ta-ta	‘to taste’
pɬt [kŋ]	→	pɬt-kŋ-kŋ-ɬp	‘tree bark’

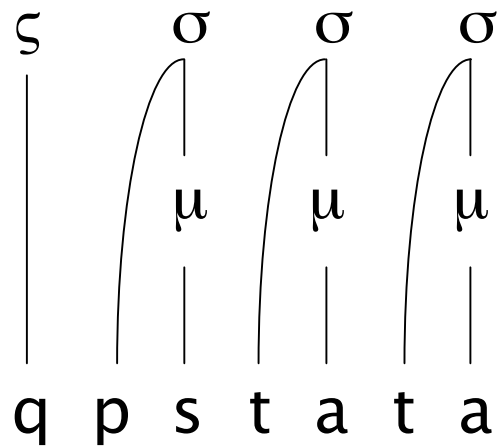
Bella Coola reduplication

Bagemihl: copying begins at the first syllable



Alternative syllabification

Fricatives can head syllables, and stray consonants are morales “semisyllables”.

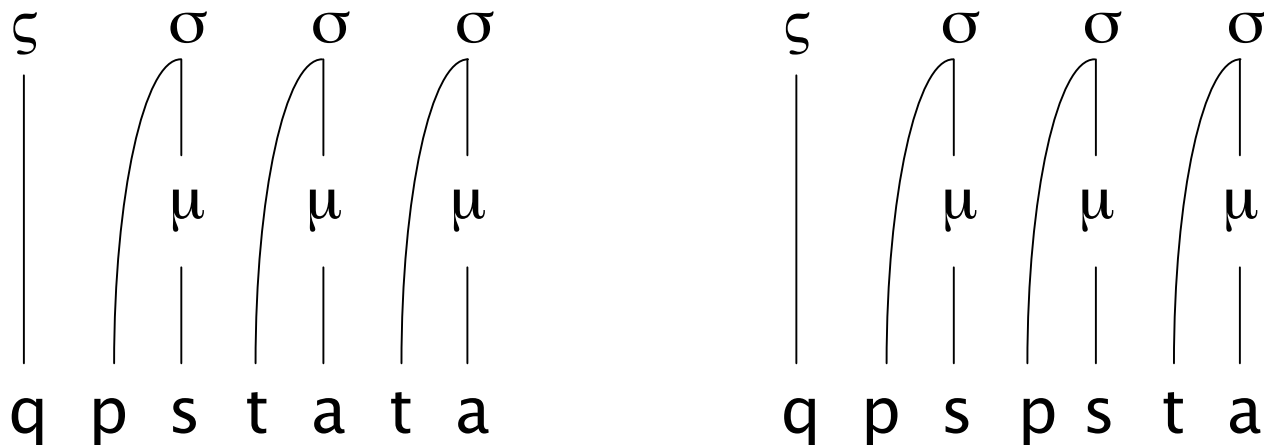


Cho and King (2003)

Unmarked Reduplicants

The Emergence of the Unmarked (TETU) in Optimality Theory prevents copying of fricative-headed syllables:

*Peak/Obstruent >> Align-Prefix

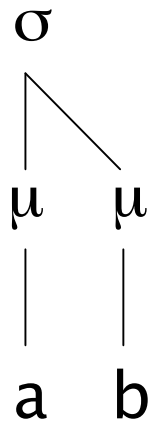


Carlson (1997)

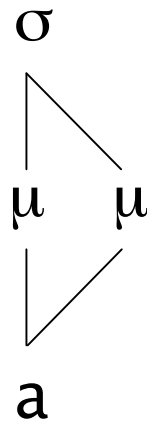
Branching moras affect timing

examples from Levantine Arabic

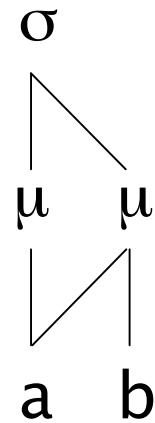
/ʕi.nab.hum/
'their grape'



/ki.taa.bi/
'my book'



/ki.taab.hum/
'their book'



Arabic consonant durations

	V_{μ} C_{μ}	V_{μ} VC_{μ}
Jordanian	88.4 ms	67.6 ms
Syrian	114.4	80.9
Lebanese	81.2	66.9

Broselow, Huffman, Chen, and Hsieh (1995)

Broselow, Chen, and Huffman (1997)

Arabic long-vowel durations

	V_{μ} V_{μ}	V_{μ} VC_{μ}
Jordanian	161.0 ms	131.6 ms
Syrian	123.9	112.2
Lebanese	114.2	97.8

Broselow, Huffman, Chen, and Hsieh (1995)

Broselow, Chen, and Huffman (1997)

Crosslinguistic V durations

	V_{μ}	$V_{\mu}VC_{\mu}$	$V_{\mu}V_{\mu}$
Arabic	1	1.6	1.9
Malayalam	1	1.6	2.2
Runyambo	1	1.4	2.2

Broselow, Chen, and Huffman (1997)
Hubbard (1994, 1995)

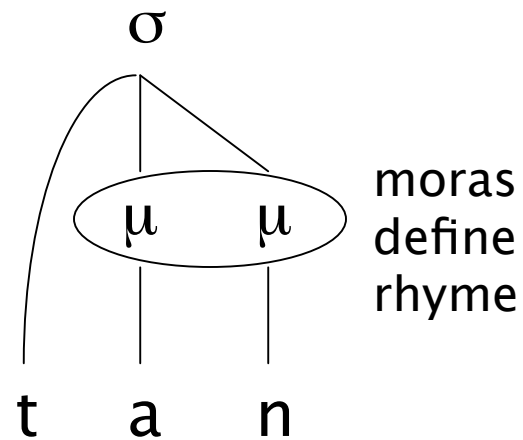
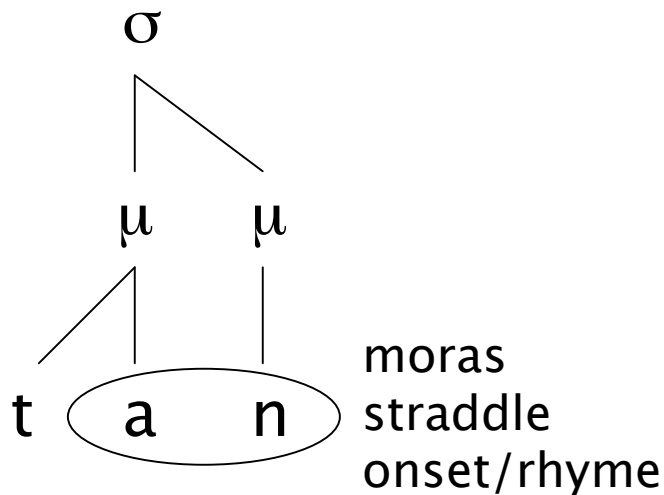
Onset/Rhyme

Zhùyīn fúhào or *bōpōmōfō* :
onset (+ glide) + rhyme

/baj/	ㄅ ㄞ	<b-ai>
/daj/	ㄉ ㄞ	<d-ai>
/daw/	ㄉ ㄠ	<d-au>
/bəŋ/	ㄅ ㄥ	<b-eng>
/bjaw/	ㄅ ㄟ ㄠ	<b-i-au>

Onset/Rhyme in Syllables

Poser includes onset/rhyme systems in his typology, but it's not clear if this is consistent with his moraic representations.



Interim Summary 3

- The evidence for CV as a moraic constituent mora is not persuasive.
- The overall consensus in the field is for onsets linked directly to the syllable node.
- The “moraic writing system” analysis is not available.

CV as a syllable

An efficient core
with supplementation

Core syllables

- Every language has CV syllables
- Some also have V syllables
- (C)V is the minimal syllable
 - “core”
 - “universal”
 - “maximally unmarked”
- Cf. Onset, NoCoda, *Complex

“Core”, not “Light”

Tagalog recent perfective:

ka-ta-trabaho

‘work’

ka-bo-bloaut

‘give special treat’

If relevant notion were light syllable, would expect CCV instead.

Japanese syllables

Basic (C)V with five deviations:

- Coda N
- Coda C as part of geminate
- Long vowel
- Complex onset
- Novel CV in borrowings

Supplemented Japanese syllables

	Spoken	Written
CVN	+ mora	+ kana
CVC	+ mora	+ small kana
CVV	+ mora	+ kana
CyV	no new mora	+ small kana
novel CV	no new mora	+ small kana

Japanese *kana* and moras

- By the nature of Japanese phonology, nearly every deviation from CV yields a new mora.
- But CyV and borrowed CV sequences are also written with two *kana*.
- No good correlation between addition of *kana* and addition of mora.

Mycenaean and Linear B

- Core CV as in typical syllabary.
- Complex onsets written with echo vowel.
 - though not /s/ before a stop
- Codas not written.
 - cf. special case of final /ks/ and sonority
- No indication of vowel or consonant length.
 - except /w/ as second half of diphthong
- Small number of “optional” signs.

Unwritten Segments in Linear B

/k ^h al.kós/	<ka-ko>	‘bronze’
/poj.měn/	<po-me>	‘shepherd’
/g ^w a.si.léws/	<qa-si-re-u>	‘chief’
/wás.tu/	<wa-tu>	‘town’
/am.phí/	<a-pi>	‘around’
/spér.ma/	<pe-ma>	‘seed’

Onset Clusters in Linear B

/knōs.sós/	<ko-no-so>	‘Knossos’
/trí.pos/	<ti-ri-po>	‘tripod’
/ktoj.nǎ.hōn/	<ko-to-na-o>	‘of lands’
/de.smōjs/	<de-so-mo>	‘with bands’
/ar.thmós/	<a-to-mo>	‘league’
/sta.thmós/	<ta-to-mo>	‘sheepfold’

Final “Onset” Clusters

/á.kso.nes/	<a-ko-so-ne>	‘axles’
/aj.ksmáns/	<ai-ka-sa-ma>	‘spear points’
/k ^w sí.p ^h e.he/	<qí-si-pe-e>	‘two swords’
/wá.naks/	<wa-na-ka>	‘king’
/aj.thí.ok ^w s/	<ai-ti-jo-qo>	‘Ethiopian’

Supplemented Linear B syllables

	Spoken	Written
CVC	+ mora	no sign
CVks	+ mora	+ sign
CV: , CVj	+ mora	no sign
CVw	+ mora	+ sign
CCV	no new mora	+ sign

Linear B signs and moras

- By the nature of Greek phonology, only some deviations from CV yield a new mora.
- The scribal practice largely ignored coda consonants and other moraic elements, but did write clusters with non-falling sonority.
 - salience of onset due to C in CV ?
- Terrible correlation between addition of sign and addition of mora.

Complex syllables in Cypriot

Similar to Linear B (and related to it), but writes most coda consonants with an echo vowel or dummy <e> word-finally.

/per.séw.tāj/	<pe-re-se-u-ta-i>	‘for Perseutas’
/dál.ton/	<ta-la-to-ne>	‘tablet’
/án.thrō.pos/	<a-to-ro-po-se>	‘man’
/gár/	<ka-re>	‘for’
/spér.ma/	<se-pe-re-ma>	‘seed’

Supplemented Cypriot syllables

	Spoken	Written
CVC	+ mora	+ sign
CVC _i .C _i	+ mora	no sign
CV:	+ mora	no sign
CVV	+ mora	+ sign
CCV	no new mora	+ sign

Cypriot signs and moras

- The essential CV system hasn't changed.
- Still deviates from a “moraic” system:
 - Nonwriting of long vowels and geminates.
 - Writing of onset clusters.
- Supplementary strategies have simply become more plentiful.
- No useful correlation between addition of sign and addition of mora.

Intentional minimality in Maya

For signs whose origin is known with some confidence, we can observe truncation of the morphographic value to CV.

<pa>	<i>paw</i>	'net'
<ka>	<i>kay</i>	'fish'
<ta>	<i>tana</i>	'house'
<po>	<i>pom</i>	'tree sap'

Complex syllables in Maya

Writes coda consonants with a “synharmonic” vowel — or if disharmonic, indicates a complex vowel with length or /h/.

/tsul/	<tsu-lu>	‘dog’
/mūt/	<mu-ti>	‘bird’
/pits/	<pi-tsi>	‘play ball’
/čihk/	<či-ku>	‘coati’
/pat/	<pa-ta>	‘acquire shape’
/bāk/	<ba-ki>	‘bone’

Supplemented Maya syllables

	Spoken	Written
CVC	no new mora?	+ sign CV _i
CV:C	+ mora	+ sign CV _k
CVhC	+ mora?	

The extra signs serve not to mark moraic information, but rather to partly indicate linguistically important distinctions: a coda consonant, and vowel length or /h/.

Syllabic signs in Akkadian

Basic

Combined

V

CV-V = CV:

CV

CV-VC = CVC

VC

CV-V-VC = CV:C

CVC

All vowels and consonants are written, but vowel length and gemination are often omitted.

Complex syllables in Akkadian

/ma.la/	<ma-la>	‘whatever’
/baš.mu/	<ba-aš-mu>	‘pertains’
/nak.la/	<nak-la>	‘clever’
/liq.tī/	<liq-ti>	‘selections’
/as.niq/	<as-niq>	‘I checked’
/ṭup.pā.ni/	<ṭup-pa-a-ni>	‘tablets’

Akkadian moraicity

/ra.bîm/

‘of the big one’

<ra-bi-im>

<ra-bi-i-im>

/i.kaš.šas.si/

‘defeats her’

<i-ka-ša-si>

<i-ka-aš-ša-si>

<i-ka-ša-as-si>

<i-ka-aš-ša-as-si>

Supplemented Akkadian syllables

	Moras	Graphs
CV	1 mora	1 in CV
CVC	2 moras	1 in CVC 2 in CV-VC
CV:	2 moras	1 in CV 2 in CV-V
CV:C	3 moras (?)	1 in CVC 2 in CV-VC 3 in CV-V-VC

Akkadian mappings

- Every sign is a possible syllable.
- Combinations of signs extend the system:
 - CV–VC for convenience or by necessity
 - CV–V for greater explicitness
- Priority to write every segment; moraic distinctions secondary.
 - relevance of segments to all languages, even if written with a syllabary

Korean syllable blocks

Alphabetic *han'gŭl* defines “syllable” on a partly morphological basis.

닭이

<talk-i>

[tal.gi]

‘chicken (SUBJ)’

값을

<kaps-ŭl>

[kap.sŭl]

‘price (OBJ)’

Similar abstraction from surface syllabification occurs in syllabaries.

The notion “syllabary”

Poser (1992):

“each syllable is represented by a distinct graph.”

Sproat (2000):

“ought to imply that every syllable of the language has a graphemic symbol associated with it.”

The notion “alphabet”

Poser (1992):

“each segment is represented by a distinct graph.”

Sproat (2000):

“symbols available to represent every phonemic segment of the language.”

Alphabets in practice

Ideally, every segment has a graph.

$[p] \rightarrow p$

$[i] \rightarrow e$

But in many cases, digraphs or other strategies are used to represent certain segments.

$[\check{c}] \rightarrow ch$

$[\eta] \rightarrow ng$

Syllabaries in practice

Ideally, every syllable has a graph.

[ta] → た

[ko] → こ

But in many cases, digraphs or other strategies are used to represent many syllables.

[tan] → たん

[kyo] → きよ

Efficiency in Japanese

Phonemes 20
(14 C, 5 V, plus N)

Syllables 276
(CVX, some CV excluded)

Kana 46
(with diacritics = 25 more)

Practical implementation

- An alphabet can reasonably reach full coverage of the segments.
 - but of course often doesn't
- A syllabary must be much larger to cover all the possible syllables of a language.
 - so this situation is rarely achieved
 - Yi, perhaps Chinese
 - for simple reasons of scale, not because the syllable isn't really the relevant unit

Implications

Adapted from Poser (1992)

1. Correspondence between Writing and Speech

“Writing systems make use of essentially the same constituents that phonologists do.”

Yes; but this set of constituents does not include the mora.

2. Phonological Constituency

“Their role in writing systems provides additional support for the syllable, mora, segment, and rhyme.”

Well, not the mora.

And we have to remember that writing systems are the products of cultural history, not pure reflections of language.

3. Alleged Atomicity of the Syllable

“The extreme rarity of syllabaries together with the evidence for subsyllabic analysis ... eviscerates the arguments for the psychological atomicity of the syllable that have been based on the putative predominance of syllabaries.”

This conclusion remains true, even if syllabaries are common; that constituent is salient and important, but not atomic.

Thanks!

Gene Buckley
Dept of Linguistics
619 Williams Hall
University of Pennsylvania
Philadelphia, PA 19104-6305

gene@ling.upenn.edu