Revised Proposal to Encode Additional Phonetic Modifier Letters in the UCS

Date: 2004-02-01

Author: Peter Constable, Microsoft

Address: One Microsoft Way
            Redmond, WA 98052
            USA

Tel: +1 425 722 1867

Email: petercon@microsoft.com

A.   Administrative

1.	Title	Revised Proposal to Encode Additional Phonetic Modifier Letters in the UCS
2.	Requester's name	SIL International (contact: Jonathan Kew), Peter Constable
3.	Requester type	Expert contribution
4.	Submission date	2004-02-01
5.	Requester's reference	L2/03-180, L2/04-045, L2/04-047
6a.	Completion	This is a complete proposal
6b.	More information to be provided?	Only as required for clarification.

B.   Technical—General

1a.	New Script? Name?	No
1b.	Addition of characters to existing block? Name?	Yes — Phonetic Extensions
2.	Number of characters in proposal	38
3.	Proposed category	A
4.	Proposed level of implementation and rationale	1 (no combining marks or jamo)
5a.	Character names included in proposal?	Yes
5b.	Character names in accordance with guidelines?	Yes

5c.	Character shapes reviewable?	Yes
6a.	Who will provide computerized font?	SIL International
6b.	Font currently available?	Yes
6c.	Font format?	TrueType
7a.	Are references (to other character sets, dictionaries, descriptive texts, etc.) provided?	Yes
7b.	Are published examples (such as samples from newspapers, magazines, or other sources) of use of proposed characters attached?	Yes
8.	Does the proposal address other aspects of character data processing?	Yes, suggested character properties are included (see section E).

C. Technical—Justification

1.	Has this proposal for addition of character(s) been submitted before?	An earlier proposal (L2/03-180) was reviewed by UTC (meeting #95), and concern was expressed at characters for which no evidence of attestation was provided. Of the 12 vowel modifier letters for which evidence was not provided in L2/03-180, evidence for 3 has since been found, and these are included in this proposal. Of the 3 consonant modifier letters for which evidence was not provided in L2/03-180, evidence for 1 has since been found, and this is included in this proposal. The remaining 11 characters for which evidence has not been found are mentioned here for information purposes, but are not included as part of the proposal.
2a.	Has contact been made to members of the user community?	Yes
2b.	With whom?	Linguists
3.	Information on the user community for the proposed characters is included?	Linguists
4.	The context of use for the proposed characters	Linguistics text books, linguistic descriptions (books, journal publications, etc.); dictionaries.
5.	Are the proposed characters in current use by the user community?	Yes
6a.	Must the proposed characters be entirely in the BMP?	Preferably

6b.	Rationale?	If possible, should be kept with other phonetic symbols in the BMP.
7.	Should the proposed characters be kept together in a contiguous range?	Preferably
8a.	Can any of the proposed characters be considered a presentation form of an existing character or character sequence?	These characters have the same appearance as superscript forms of existing characters.
8b.	Rationale for inclusion?	Like other modifier letters used for phonetic transcription, these need to be considered distinct characters and not presentation forms of existing characters.
9a.	Can any of the proposed characters be considered to be similar (in appearance or function) to an existing character?	No
9b.	Rationale for inclusion?	n/a
10.	Does the proposal include the use of combining characters and/or use of composite sequences?	No.
11.	Does the proposal contain characters with any special properties?	No.

D.   SC2/WG2 Administrative

1.   Relevant SC2/WG2 document numbers

2.   Status (list of meeting number and corresponding action or disposition)

3.   Additional contact to user communities, liaison organizations, etc.

4.   Assigned category and assigned priority/time frame

Other comments

E.   Proposed Characters

A code chart and list of character names are shown on a new page.

E.1   Code Chart

	xx0	xx1	xx2	xx3
0	ɒ	ɪ	ʂ	θ
1		ɫ	ʃ	
2	c	ɟ	ʈ	
3	ç	ɭ	ʉ	
4	ð	ɭ	ʊ	
5		ʟ	ᵾ	
6	ʒ	ɱ		
7		ɰ	ʋ	
8	f	ɲ	ʌ	
9	ɟ	ŋ		
A	ɡ	ɴ		
B			ʐ	
C	ɥ	ɵ	ʐ̢	
D			ʑ	
E	ɨ		ʒ	
F	ɩ	ɸ	ʔ	

E.2   Character Names

xx00   MODIFIER LETTER SMALL TURNED ALPHA
≈ <super> 0252

xx01   (this position left empty)

xx02   MODIFIER LETTER SMALL C
≈ <super> 0063

xx03   MODIFIER LETTER SMALL C WITH CURL
≈ <super> 0255

xx04   MODIFIER LETTER SMALL ETH
≈ <super> 00F0

xx05   (this position left empty)

xx06   MODIFIER LETTER SMALL REVERSED OPEN E
≈ <super> 025C

xx07   (this position left empty)

xx08   MODIFIER LETTER SMALL F
≈ <super> 0066

xx09   MODIFIER LETTER SMALL DOTLESS J WITH STROKE
≈ <super> 025F

xx0A   MODIFIER LETTER SMALL SCRIPT G
≈ <super> 0261

xx0B   (this position left empty)

xx0C   MODIFIER LETTER SMALL TURNED H
≈ <super> 0265

xx0D   (this position left empty)

xx0E   MODIFIER LETTER SMALL I WITH STROKE
≈ <super> 0268

xx0F   MODIFIER LETTER SMALL IOTA
≈ <super> 0269

xx10   MODIFIER LETTER SMALL CAPITAL I
≈ <super> 026A

xx11   MODIFIER LETTER SMALL CAPITAL I WITH STROKE
≈ <super> LATIN SMALL CAPITAL I WITH STROKE (see separate proposal, Constable 2004b = L2/04-047)

xx12   MODIFIER LETTER SMALL J WITH CROSSED-TAIL
≈ <super> 029D

xx13   MODIFIER LETTER SMALL L WITH RETROFLEX HOOK
≈ <super> 026D

xx14   MODIFIER LETTER SMALL L WITH PALATAL HOOK
≈ <super> LATIN SMALL LETTER L WITH PALATAL HOOK (see separate proposal, Constable 2004a = L2/04-045)

xx15   MODIFIER LETTER SMALL CAPITAL L
≈ <super> 029F

xx16   MODIFIER LETTER SMALL M WITH HOOK
≈ <super> 0271

xx17   MODIFIER LETTER SMALL TURNED M WITH LONG LEG
≈ <super> 0270

xx18   MODIFIER LETTER SMALL N WITH LEFT HOOK
≈ <super> 0272

xx19	MODIFIER LETTER SMALL N WITH RETROFLEX HOOK ≈ <super> 0273	xx25	MODIFIER LETTER SMALL CAPITAL U ≈ <super> 1D1C
xx1A	MODIFIER LETTER SMALL CAPITAL N ≈ <super> 0274	xx26	(this position left empty)
		xx27	MODIFIER LETTER SMALL V WITH HOOK ≈ <super> 028B
xx1B	(this position left empty)		
xx1C	MODIFIER LETTER SMALL BARRED O ≈ <super> 0275	xx28	MODIFIER LETTER SMALL TURNED V ≈ <super> 028C
xx1D	(this position left empty)	xx29	(this position left empty)
xx1E	(this position left empty)	xx2A	(this position left empty)
xx1F	MODIFIER LETTER SMALL PHI ≈ <super> 0278	xx2B	MODIFIER LETTER SMALL Z ≈ <super> 007A
xx20	MODIFIER LETTER SMALL S WITH HOOK ≈ <super> 0282	xx2C	MODIFIER LETTER SMALL Z WITH RETROFLEX HOOK ≈ <super> 0290
xx21	MODIFIER LETTER SMALL ESH ≈ <super> 0283	xx2D	MODIFIER LETTER SMALL Z WITH CURL ≈ <super> 0291
xx22	MODIFIER LETTER SMALL T WITH PALATAL HOOK ≈ <super> 01AB	xx2E	MODIFIER LETTER SMALL EZH ≈ <super> 0292
xx23	MODIFIER LETTER SMALL U BAR ≈ <super> 0289	xx2F	MODIFIER LETTER SMALL GLOTTAL STOP ≈ <super> 0294
xx24	MODIFIER LETTER SMALL UPSILON ≈ <super> 028A	xx30	MODIFIER LETTER SMALL THETA ≈ <super> 03B8

E.3   Unicode Character Properties

All of the proposed characters should have a general category of Lm. Compatibility decompositions should be as shown above. Other properties should match those of similar characters, such as U+02E1 MODIFIER LETTER SMALL L.

Note that the proposed compatibility decomposition for xx04 MODIFIER LETTER SMALL REVERSED OPEN E is <super> 025C. The existing character U+1D4C MODIFIER LETTER SMALL TURNED OPEN E currently has a compatibility decomposition mapping of <super> 025C. It is suggested that the latter mapping is in error, and proposed that U+1D4C should rather have a compatibility decomposition mapping of <super> 1D08.

Note also that the proposed compatibility decomposition for xx12 MODIFIER LETTER SMALL L WITH PALATAL HOOK is to a character that is not yet encoded, but has been proposed for encoding in a separate document, submitted at the same time (Constable 2003).

F.   Other Information

In general, modifier letters are used in phonetic transcription to represent secondary aspects of articulation. Secondary articulations may involve aspects of simultaneous articulation that are considered to be in some sense less dominant to the basic sound (for instance, nasalized vowels are typically conceived in terms of their oral counterparts but with the additional secondary articulation of nasalization); or they may involve a transitional articulation of a type that might otherwise be considered a complete speech sound in its own right but for various reasons is

interpreted by the linguist as a secondary element in a complex speech sound (for instance, diphthongs, or nasal onset of oral stop consonants). In some situations, the recommended transcription using the International Phonetic Alphabet would not involve a modifier letter; thus, many of the proposed characters are not officially-approved IPA notation. Nevertheless, the use of these modifier letters is fairly commonplace among linguists, even those that advocate the use of IPA.

The proposed modifier letters are of two types: those used in phonetic transcription to represent vowel-like sounds, and those used to represent consonantal sounds. These two groups will be discussed separately.

F.1   Vowel modifier letters

Vowel modifier letters are often used by linguists in transcribing diphthongs. Diphthongs are speech sounds involving two distinct but sequentially-contiguous vocalic gestures — two vowel targets. For instance, whereas the Spanish phoneme /e/ is typically spoken with a single vowel target, [e], the English phoneme /e/ is very often spoken with two vowel targets, [e] and [i]. Following the conventions of IPA strictly, the English phoneme could be transcribed as [ei] or [e͡i]. Occasionally, though, linguists will transcribe such a diphthong as [eⁱ] or [ᵉi], according to which component is considered to be secondary — an "on-glide" or an "off-glide":

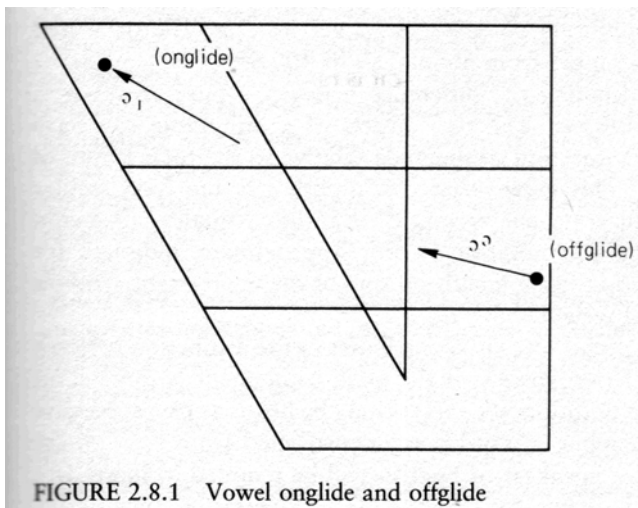FIGURE 2.8.1   Vowel onglide and offglide

Figure 1. Vowel modifier letters used to indicate "on-glide" or "off-glide" diphthongs (Clark and Yallop 1995, p. 35).

Vowel modifier letters are also sometimes used to transcribe syllables that have a marginally-vocalic nucleus or a vocalic nucleus of very short duration, such that the vowel component of the syllable seems suppressed in relation to the consonantal components.

There are already a number of vowel modifier letters encoded in the UCS. Most of these were added in Unicode 4.0 and are in the Phonetic Extensions block:

1D43	ᵃ	MODIFIER LETTER SMALL A
1D44	ᵄ	MODIFIER LETTER SMALL TURNED A
1D45	ᵅ	MODIFIER LETTER SMALL ALPHA
1D46	ᵆ	MODIFIER LETTER SMALL TURNED AE
1D49	ᵉ	MODIFIER LETTER SMALL E
1D4A	ᵊ	MODIFIER LETTER SMALL SCHWA
1D4B	ᵋ	MODIFIER LETTER SMALL OPEN E
1D4C	ᵌ	MODIFIER LETTER SMALL TURNED OPEN E
1D4E	ᵎ	MODIFIER LETTER SMALL TURNED I
1D52	ᵒ	MODIFIER LETTER SMALL O
1D53	ᵓ	MODIFIER LETTER SMALL OPEN O
1D54	ᵔ	MODIFIER LETTER SMALL TOP HALF O
1D55	ᵕ	MODIFIER LETTER SMALL BOTTOM HALF O
1D58	ᵘ	MODIFIER LETTER SMALL U
1D59	ᵙ	MODIFIER LETTER SMALL SIDEWAYS U
1D5A	ᵚ	MODIFIER LETTER SMALL TURNED M
2071	ⁱ	SUPERSCRIPT LATIN SMALL LETTER I

Table 1. Vowel modifer letters encoded in Unicode 4.0

This covers those vowel sounds that are most commonly encountered in the world's languages. This list does not include all vowel symbols used in phonetic transcription, however. In principle, any vowel gesture may potentially be one of the targets in a diphthong. Ladefoged and Maddieson (1996, p. 322) comment, "The kinds of vowels that occur as targets in diphthongs are no different from those that occur as single vowels."

When combined with modifier letters already encoded in Unicode 4.0, the vowel modifier letters proposed here cover most of the vowel symbols from the IPA and Americanist traditions:[1]

---

[1] While IPA is increasingly prevalent, the Americanist tradition is still in use, and the use of superscripts to transcribe diphthongs may be more prevalent among those that use Americanist conventions. Some vowels in the Americanist system use diacritics, but it is assumed that combining marks can be used in sequences with modifier letters as well as with other letters. Capital vowel letters are used by some in the Americanist tradition to transcribe voiceless vocoids, but this proposal does not include modifier-letter counterparts to Latin capital vowel letters. I am not aware at the present time of a user need for capital vowel modifier letters in order to transcribe a voiceless, secondary component of a diphthong using Americanist conventions.
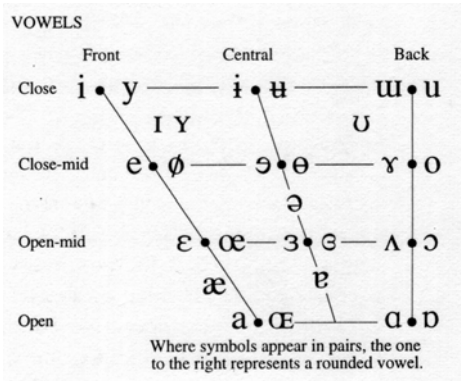
Figure 2. IPA vowels (IPA 1999, p. ix).

American Usage Vowel Symbols

		Front		Central		Back	
		Unround	Round	Unround	Round	Unround	Round
High	(Higher)	i	ü	ɨ	ʉ	ï	u
	Lower	I	Ü	ɪ	ʊ	Ï	U
Mid	Higher	e	ö		ə	ë	o
	Lower	ɛ	ɔ̈		ʌ		ɔ
Low		æ			a/ɑ		
Lower-Low		a		ɑ	ɒ		

Figure 3. Americanist vowels (Pullum and Ladusaw 1996, p. 298).[2]

The vowel portion of the overall proposal is summarized in Table 2:

---

[2] There is some variation within Americanist usage. Whereas Pullum and Ladusaw show a small capital I for the front unrounded lower-high vocoid, many represent this vocoid using small iota. Also, some use a small v with hook for the back round lower-high vocoid, rather than the small capital u shown here. Barred iota and barred v-hook for central lower-high vowels are not used, however.

Character			Samples
xx00	ᴅ	MODIFIER LETTER SMALL TURNED ALPHA	Figure 17
xx06	ɜ	MODIFIER LETTER SMALL REVERSED OPEN E	Figure 9
xx0E	ɨ	MODIFIER LETTER SMALL I WITH STROKE	Figure 8, Figure 12
xx0F	ɩ	MODIFIER LETTER SMALL IOTA	Figure 16
xx10	ɪ	MODIFIER LETTER SMALL CAPITAL I	Figure 9, Figure 15
xx11	ɪ̵	MODIFIER LETTER SMALL CAPITAL I WITH STROKE	Figure 14
xx1C	ɵ	MODIFIER LETTER SMALL BARRED O	Figure 9
xx23	ʉ	MODIFIER LETTER SMALL U BAR	Figure 8
xx24	ʊ	MODIFIER LETTER SMALL UPSILON	Figure 11
xx25	ᴜ	MODIFIER LETTER SMALL CAPITAL U	Figure 16
xx28	ʌ	MODIFIER LETTER SMALL TURNED V	Figure 6, Figure 12

Table 2. Proposed vowel modifer letters

The following samples serve to illustrate the use of vowel modifier letters in general to transcribe diphthongs, and also to demonstrate attestation of the vowel modifier letters proposed. The samples will contain vowel modifier letters that are already encoded as well as those being proposed; those already encoded will be highlighted in blue; those being proposed, in red.

Figure 4. Vowel modifier letters: a, schwa, u (Czaykowska-Higgins and Willett 1997, p. 408).

Figure 5. Vowel modifier letter: schwa (Bessell 1998, p. 5).

Panjabi (Central-Southern)

[ləˈkɾɪ̀ðɑːˈrerɑ̤] 'a wooden cart'

[kˣʌˈtɛ̪ðɪ̀kʰˠˈtɑe̪ˀ] 'the sourness of the sour-fruit'

[ˌɡʷfi̤ʷ ʊ̯mbə̯ɾˈɡfiɛ̯ːr̥] 'whirlpool'

Figure 6. Vowel modifier letter: turned v (Laver 1994, p. 560).

3. *Kāna mīyaᵉaᵘ ('aaᵘ) päᵉjtk mạnitō pạ=ma'kạminạnk iᵘ kitcikutäminān*
   (12:14–15)

   'Doubtless it is one of the manitous **that has come to take away** this fire of ours'.

4. *Mīsa iᵋíᵘ pi=nạndōpạnitawiyạn?* (18:7)

   'And so **you have come looking for me?**'

Figure 7. Vowel modifier letters: e, open-e, u (Malone 1999, p. 353).

a - a	ádzà	+ akpà	→	ádzãkpà	a clever man
i - a	opi	+ àyi	→	opⁱayi	Lendu woman
ɨ - a	ìndrì	+ akpà	→	ìndrᵻakpà	male goat
e - a	àwè	+ àyi	→	àwᵉayi	pygmy woman
ɛ - a	ɨbhɛ̀	+ akpà	→	ɨbhᵋakpà	big fish
u - a	ìndrũ	+ àyi	→	ìndrᵘayi	Ngiti woman
ʉ - a	à'ʉ̈	+ akpà	→	à'ᵚakpà	cock
o - a	abvo	+ àyi	→	abvᵒayi	widow
ɔ - a	tɨ̀tɔ̀	+ akpà	→	tɨtᵓakpà	liar

Figure 8. Vowel modifier letters: i, i-bar, e, open-e, u, u-bar, o, open-o (Lojenga 1994, p. 90).

[fi̤ʷᵊ̯ɹ̝ʷbi̝tɒvʷɲʷḭː] 'cemetery-like'

[ˈbɱɛ̃ᵑɳ̊kɑ] 'female citizen of Brno'

[ˈrɑɾɑːʃʷɛ̞k] 'little devil'

Danish

[kʷʰø̞ᴸpmhã̤ʊ̃ʔʷn̥] 'Copenhagen'

[ø̞sʷtʷ ə̯hʷʊ̯ʔʷlʲʷ] 'cheese-slicer'

Dutch (Amsterdam)

[kʰsəˀɪ̯] 'calf'

[kˣʊutʷ] 'cold'

Finnish (Kouvola)

[pɑːlʲø̞svɑːtɛ̠] 'cloak'

[mʲʷ ø̞n̪tæ] 'to admit'

Scottish Gaelic (Skye)

[kʲaʎiçː] 'of an old woman'

[kʰɑ̤ːɬᵌ̯ɪ̯ʃ] 'accompanying him/along with him'

Figure 9. Vowel modifier letters: o-bar, small capital I, schwa, reversed open e (Laver 1994, p. 559).

(tw° nettoyere)
(tɽ Patric)
(tj gytje)
tr ytre
-
(kw° Equador)

Figure 10. Vowel modifier letter: open-o (Brink et al 1998, p. 99).

(3) *Variants*—Refined RP unrounds, raises, and centralizes the starting point, that we have [ɐ̈]. This produces a cluster of unrounded back open vowels or diphthongs: /ɑː/=[ɑː], /ɜː/=[ɑː], /aɪ/=[ɐ̈ɪ], /ɔɪ/=[ɐ̈ɪ] and /aʊ/=[ɐ̈ʊ], the cumulative effect of which is the so-called 'plummy' effect associated with this accent.

Figure 11. Vowel modifier letter: upsilon (Cruttenden 2001, p. 133).

		First element		Second element	
Breathy	**Vowels**	F1	F2	F1	F2
ព្រាប	[iːᵉ]	365	2556	531	2115
ទៀន	[iːᵉ]	326	2658	548	2149
ជើ	[iːᴧ]	364	1587	508	1544
ទុត	[uːᴧ]	401	937	554	1284
ផ្ទេ	[ɛⁱ]	535	2280	329	2551
ភ្លៅ	[əᵘ]	580	1336	428	882
Clear	**Vowels**	F1	F2	F1	F2
ត្រចៀក	[iːᵉ]	340	2784	571	2275
ប្រៀង	[iːᵊ]	478	1416	542	1453
ធុត	[uːᴧ]	414	1027	547	1304
ក	[ɛⁱ]	620	1943	341	2829
ជ្រៅ	[əᵘ]	557	1313	452	840
មុន	[əːⁱ]	566	1519	369	1518
ត្រី	[əːⁱ]	526	1577	378	1447
កោស	[ɔːᵒ]	713	1121	615	979
កើត	[aːᵊ]	768	1489	554	1496
ថៃ	[aⁱ]	792	1983	436	2659
ចៅ	[aᵘ]	855	1626	510	880

Figure 12. Vowel modifier letters: open-e, schwa, i, i-bar, o, u, turned-v, (Wayland and Allard 2001, p. 76).

listed in the earlier table), the initial voiced stops are unexploded. These are not sequences of the form **dᵊt**, but are simply homorganic pairs of stops, with the first member being voiced and unreleased, and the second being voiceless and, on some occasions, also ejective and affricated.

Figure 13. Vowel modifier letter: schwa (Ladefoged and Maddieson 1996, p. 80).

Figure 14. Vowel modifier letter: small capital i-bar (Bailey 1985, p. xxv).

Figure 15. Vowel modifier letter: small capital i (Bailey 1985, p. xxvi).

Figure 16. Vowel modifier letters: iota, small capital u (Floyd 1981, p. 19).

Secondary Articulation	Description in terms of tongue and lip position	Symbolization	
		simul.	off-gl.
Palatalization	[i] or [y]		t^y
Labialization	[u] or [w]	ţ	t^w
Palato-labialization	[ʉ]		$t^{ʉ}$
Velarization	[ï]	ŧ	$t^ï$
Pharyngealization	tongue root towards pharynx wall	ţ	$t^ᴅ$

Figure 17. Vowel modifier letter: small turned alpha (Floyd 1981, p. 105).

The vowel modifier letters in the following table are those that would be needed to provide complete coverage for IPA and Americanist vowel symbols as shown in Figure 2 and Figure 3 but for which attestation is not yet documented. Code positions as shown have been left empty in the code chart in §E.1 to allow for the possibility that attestation for one or more of these might be found in the future.

æ	MODIFIER LETTER SMALL AE	xx01
ɘ	MODIFIER LETTER SMALL REVERSED E	xx05
ɞ	MODIFIER LETTER SMALL CLOSED REVERSED OPEN E	xx07
ɤ	MODIFIER LETTER SMALL RAMS HORN	xx0B
ø	MODIFIER LETTER SMALL O WITH STROKE	xx1B
œ	MODIFIER LETTER SMALL LIGATURE OE	xx1D
Œ	MODIFIER LETTER SMALL CAPITAL OE	xx1E
ʉ	MODIFIER LETTER SMALL CAPITAL U WITH STROKE	xx26
ʏ	MODIFIER LETTER SMALL CAPITAL Y	xx2A

Table 3. Vowel modifer letters not currently proposed for encoding

F.2   Consonant modifier letters

Consonant modifier letters are often used to transcribe articulatory modifications that may apply to a wide variety of consonantal sounds, such as aspiration (typically transcribed as [ʰ]) or labialization (typically transcribed as [ʷ]). Consonantal modifier letters can also be used to transcribe sounds that involve a secondary consonantal articulation in addition to the dominant consonant, either simultaneously or as a transitional effect, such as a lateral release (typically transcribed as [ˡ]).

The most commonly-used consonant modifier letters are already encoded in the UCS. Several others are also in use, however. The inventory that seems to be needed includes nasals (e.g. to transcribe nasal onset or release of oral stops), fricatives (for fricative release of stops), approximants and some stops. Modifier counterparts for other symbols, such as clicks and trills, are not required. The samples shown below demonstrate attestation of most of the proposed inventory. The proposed consonant modifiers are listed along with an index to the samples illustrating each one in Table 4 to Table 7.

Note that a modifier counterpart to small c is proposed. The small letter c is used to represent a palatal stop. In fact, the modifier that is attested (see Figure 24) is c-cedilla, which represents a palatal fricative. It is assumed that that a voiceless affricate with a secondary palatal fricative component can be represented using a sequence < modifier letter small c, combining cedilla >. This requires, though, that the modifier letter small c be encoded.

Note also that modifier letters l-palatal hook and t-palatal hook are proposed. While the use of palatal hook for indicating palatalization is no longer an IPA recommendation, l-palatal hook and t-palatal hook are proposed here because they are attested, as seen in the samples.

Character			Samples
xx16	ᵐ	MODIFIER LETTER SMALL M WITH HOOK	Figure 21
xx18	ᶮ	MODIFIER LETTER SMALL N WITH LEFT HOOK	Figure 18, Figure 19, Figure 20, Figure 22, Figure 23
xx19	ᶯ	MODIFIER LETTER SMALL N WITH RETROFLEX HOOK	Figure 19, Figure 20
xx1A	ᴺ	MODIFIER LETTER SMALL CAPITAL N	Figure 19, Figure 20

Table 4. Proposed nasal consonant modifier letters and figures that illustrate them

Character			Samples
xx02	ᶜ	MODIFIER LETTER SMALL C (base for c-cedilla)	Figure 24, Figure 28
xx03	ᶝ	MODIFIER LETTER SMALL C WITH CURL	Figure 29, Figure 33
xx04	ᶞ	MODIFIER LETTER SMALL ETH	Figure 25
xx08	ᶠ	MODIFIER LETTER SMALL F	Figure 24, Figure 26, Figure 34
xx12	ᶨ	MODIFIER LETTER SMALL J WITH CROSSED-TAIL	Figure 25
xx1F	ᶲ	MODIFIER LETTER SMALL PHI	Figure 24
xx20	ᶳ	MODIFIER LETTER SMALL S WITH HOOK	Figure 18, Figure 24
xx21	ᶴ	MODIFIER LETTER SMALL ESH	Figure 24, Figure 27, Figure 28, Figure 33, Figure 34
xx2B	ᶻ	MODIFIER LETTER SMALL Z	Figure 25, Figure 27, Figure 33, Figure 34
xx2C	ᶼ	MODIFIER LETTER SMALL Z WITH RETROFLEX HOOK	Figure 18, Figure 25
xx2D	ᶽ	MODIFIER LETTER SMALL Z WITH CURL	Figure 18, Figure 29, Figure 33
xx2E	ᶾ	MODIFIER LETTER SMALL EZH	Figure 25, Figure 28, Figure 34
xx30	ᶿ	MODIFIER LETTER SMALL THETA	Figure 24, Figure 27, Figure 30, Figure 31, Figure 32

Table 5. Proposed fricative consonant modifier letters and figures that illustrate them

The only fricative symbol for which attestation of a corresponding modifier letter was not found is small h with stroke. Code position xx0D has been left empty in the code chart in §E.1 to allow for the possibility that attestation of this modifier letter might be found in the future.

Character			Samples
xx0C	ꞯ	MODIFIER LETTER SMALL TURNED H	Figure 35
xx13	ˡ	MODIFIER LETTER SMALL L WITH RETROFLEX HOOK	Figure 40
xx14	ᶘ	MODIFIER LETTER SMALL L WITH PALATAL HOOK	Figure 40
xx15	ᴸ	MODIFIER LETTER SMALL CAPITAL L	Figure 18
xx17	ꟺ	MODIFIER LETTER SMALL TURNED M WITH LONG LEG	Figure 42
xx27	ᶹ	MODIFIER LETTER SMALL V WITH HOOK	Figure 39

Table 6. Proposed approximant consonant modifier letters and figures that illustrate them

The only approximant symbol for which attestation of a corresponding modifier letter was not found is small turned y. Code position xx29 has been left empty in the code chart in §E.1 to allow for the possibility that attestation of this modifier letter might be found in the future.

Character			Samples
xx09	ᶡ	MODIFIER LETTER SMALL DOTLESS J WITH STROKE	Figure 29
xx0A	ᶢ	MODIFIER LETTER SMALL SCRIPT G	Figure 29, Figure 38
xx22	ᶵ	MODIFIER LETTER SMALL T WITH PALATAL HOOK	Figure 41
xx2F	ˀ	MODIFIER LETTER SMALL GLOTTAL STOP	Figure 36, Figure 37

Table 7. Proposed stop consonant modifier letters and figures that illustrate them

In the samples below, modifiers that are already encoded will be highlighted in blue, while those being proposed will be highlighted in red.
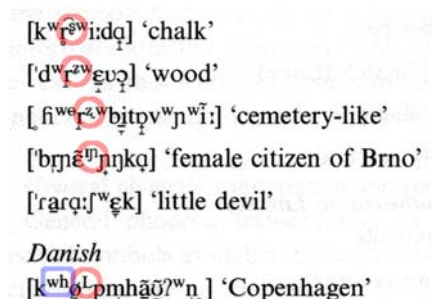
[kʷɾˢʷiːdɑ] 'chalk'

[dʷɾᶻʷɛʊɔ] 'wood'

[ɦʷᵉɾᶻʷbitɒvʷɲʷĩː] 'cemetery-like'

[ˈbɾɲɛ̃ᵐɳkɑ] 'female citizen of Brno'

[ˈɾaɾɑːʃʷɛk] 'little devil'

Danish
[kʷʰøᴸpmhãɵ̃ʔʷn̩] 'Copenhagen'

Figure 18. Consonant modifier letters: s-hook, z-curl, z-retroflex hook, n-lefthook, small capital L (Laver 1994, p. 559).

Labial	Dental	Dent/Al	Alveol	Post-al	R'flex	Palatal	Velar	Uvular	**Total**
[ᵐb	ⁿd	ⁿd ~ ⁿd	ⁿd	ⁿd	ⁿɖ	ⁿɟ	ⁿg	ᶰɢ]	

Figure 19. Consonant modifier letters: n-left hook, n-retroflex hook, small capital n (Laver 1994, p. 583).

Labial	Dental	Dent/Al	Alveol	Post-al	R'flex	Palatal	Velar	Uvular	**Total**
[bᵐ	dⁿ	dⁿ ~ dⁿ	dⁿ	dⁿ	ɖⁿ	ɟⁿ	gⁿ	ɢᴺ]	

Figure 20. Consonant modifier letters: n-left hook, n-retroflex hook, small capital n (Laver 1994, p. 584).

(151) [ᵐbu]   'ear of corn'
[ᵐpʰɛt]   'broom'
[ᵐfwaɾm]   'head cold'
[ᵐvʷaŋ]   'circumcision'

[ⁿdam]   'stick'
[ⁿtʰami]   'this year'
[ⁿsiɾ]   'old'

[ⁿdʒaŋ]   'bow'
[ⁿtʃʰon]   'fingernail'
[ⁿʃi]   'bee'
[ⁿʒi]   'goblin'

[ŋgas]   'Angas'
[ŋkʰi]   'liver'
[ŋɣik]   'rock'

Figure 21. Consonant modifier letters: m-hook, n-left hook (Burquist 2001, p. 118).

sequence (ibid., p. 16). Sequences spanning a morpheme boundary furnish many examples of two-phoneme sequences. The difference is shown by the words ki.ʈa.ⁿʈi "circumcised boy" and kar.kaŋ-.ʈi "kitehawk-erg."

Figure 22. Consonant modifier letter: n with left hook (Evans 1995, p. 732).

(52)   root   2nd person
a. puht   ᵐbʸuhtu   'you went out'
tih   ⁿdʸihu   'you arrived'
ciŋ   ⁿjiɲu   'you bathed'
ken   ŋgʸenu   'you looked'

Figure 23. Consonant modifier letter: n with left hook (Pigott 1997, p. 469).

Labial	Lab-dnt	Dental	Dent/Al	Alveol	Post-al	R'flex	Palatal	Velar	Uvular	Total
[pᶲ	pᶠ	tᶿ	t̪ᶿ ~ tˢ	tˢ	tʃ	ʈˢ	cᶜ	kˣ	qᵡ]	

Figure 24. Consonant modifier letters: phi, f, theta, esh, s-hook, c-cedilla (Laver 1994, p. 581).

Labial	Lab-dnt	Dental	Dent/Al	Alveol	Post-al	R'flex	Palatal	Velar	Uvular	Total
[bᵝ	bᵛ	d�ð	d̪ð ~ dᶻ	dᶻ	dʒ	ɖᶼ	ɟʲ	gᵞ	ɢᵡ]	

Figure 25. Consonant modifier letters: eth, z, ezh, z-retroflex hook, crossed-tail j (Laver 1994, p. 581).

(19) *Distribution of Old Allemanic non-fricative phonemes*

	initial			medial			final	
p	t	k	p	t	k	p	t	k
			pp	tt	kk	pp	tt	kk
pᶠ	tˢ	kˣ	pᶠ	tˢ	kˣ	pᶠ	tˢ	kˣ

Figure 26. Consonant modifier letter: f (Kraehenmann 2001, p. 139).

[nˡ³d̪ θɾʌxˑk] '(family) washing'

Scottish Gaelic (Islay)

[kʰʷθ̪ʉˣʷt̪ʷᵖ] 'cat'

[n̪d̪ᵏɑʔə̥] 'the day'

Figure 27. Consonant modifier letters: theta, esh, z (Laver 1994, p. 559).

Affrication is transcribed in any of three ways. One way is to write a small superscript homorganic fricative symbol after the stop symbol, as in [tˢ, tʃ, dᶾ, cˢ, ɟʲ, kˣ, gˠ]. The alternative modes of transcription involve writing both components on the line and either joining them with a linker diacritic, as in [t͡s], or joining the two symbols physically together, as in [ts].

Figure 28. Consonant modifier letters: esh, ezh, c-cedilla (Laver 1994, p. 364).

otherwise. However, the conventional way of transcribing preploded nasals, i.e. a nasal preceded by its homorganic stop [ᵇm, ᵈn, ᶨɲ, ᵍŋ], has been retained in the present material.

    Second, as also acknowledged by Bishop (1996:235) for Kensiw, the preploded nasals are historically and cognitively developments from simple nasals and have simple nasal reflexes in other Mon-Khmer languages. Importantly, reduplications of preploded nasals are always realised as the simple nasal counterpart: [səmsɔᵇm] /smsɔm/ 'to buzz around a nest', [hənhəᵈn] /hnhən/ 'to devour', [jɲjɛᶨɲ] /jɲjɛɲ/ 'to dream', [jᶻənjᶻɛᵍŋ] /jɲjɛɲ/ 'wide'. Furthermore, Malay loanwords which originally have final nasals are usually realised with the preploded counterpart: [hajaᵇm] from Malay *ayam* 'poultry', [bulaᵈn] from Malay *bulan* 'moon', [kucᶜᵉᵍŋ] from Malay *kucing* 'cat'.

Figure 29. Consonant modifier letters: c-curl, script g, dotless j-stroke, z-curl (Burenhult 2001, p. 35).

25.   pənəsθət   ʔə   tᶿəyʔ   ċəlʔċs   tᶿə   sceełtən.
       sprinkle self   with that   its scale the salmon

26.   niʔ xʷəstəlʔiʔtiyamʔ   tᶿəyʔ, ste   ʔəwʔniis
       they were stuck on   that   be like
       The little things would stick on, just like a garment.

Figure 30. Consonant modifier letter: theta (Hukari et al, p. 43).

ʥ	D-Curly-tail-Z ligature	Voiced alveolo-palatal affricate *Superseded by 104+183*	216	02A5	E2FB
θ	Superscript theta	Voiceless dental fricative release	217	------	E21B
ə	Superscript schwa	Mid central vowel release	218	------	E21A

Figure 31. Consonant modifier letter: theta (IPA 1999, p. 179).

(13) a. *ni*   *qʷəl-ət-əs*   θə   *słeniʔ*   tᶿə   *sceːłtən*
          AUX cook-TR-3ERG   DET   woman   DET   salmon
          'The woman cooked the salmon.'

      b. *ni*   *qʷəl-əm*   θə   *słeniʔ*   ʔə   tᶿə   *sceːłtən*
          AUX cook-INTR DET   woman   OBL   DET   salmon
          'The woman cooked the salmon.'

Figure 32. Consonant modifier letter: theta (Gerdts 1998, p. 309).

and Polish. As shown in (3), the stem-final consonants /t d/ in Polish (Rubach 1984) are affricated into the alveolo-palatals [tᶝ dᶻ] when followed across a morpheme boundary by the locative singular /ɛ/, verbalising /e i/ or the feminine suffix /itᶝ+a/, by virtue of Coronal Palatalisation.

(3) *Polish Coronal Palatalisation* (Rubach 1984)

	nom sg				
a.	brat	'brother'	*loc sg* /ɛ/	brac+ie	[bratᶜ+ɛ]
	cud	'miracle'		cudz+ie	[cudᶻ+ɛ]
b.	lot	'flight'	*verbalising* /e/	lec+ie+ć	[letᶜ+e+tᶜ]
c.	brud	'dirt'	*verbalising* /i/	brudz+i+ć	[brudᶻ+i+tᶜ]
d.	kot	'cat'	*fem* /itᶜ+a/	koc+ic+a	[kotᶜ+itᶜ+a]

Another source of sibilant affricates is Strident Assimilation, as in Polish (Rubach 1994). The anterior obstruents /t d/ in Polish are optionally affricated before sibilant fricatives or affricates within a lexical item or across word boundaries, as shown in (4).[4]

(4) *Polish Strident Assimilation* (Rubach 1994)

od soboty	[t s] ~ [tˢ s]	'since Saturday'
twardszy	[t ʃ] ~ [tʃ ʃ]	'harder'
odcedzić	[t tˢ] ~ [tˢ tˢ]	'drain'
świadczyć	[t tʃ] ~ [tʃ tʃ]	'witness'
odznaczyć	[d z] ~ [dᶻ z]	'distinguish'
budżet	[d ʒ] ~ [dᶾ ʒ]	'budget'
przed dzwonkiem	[d dᶻ] ~ [dᶻ dᶻ]	'before the bell'
oddzielić	[d dᶾ] ~ [dᶾ dᶾ]	'separate'

Figure 33. Consonant modifier letters: c-curl, esh, z, z-curl (Kim 2001, p. 93).

Similarly, there is a sound change from Proto-Bantu to Mvumbo which also shows plosive assibilation before the high vowels /i/ and /u/. As shown in (14a), the plosives /b d t g k/ in Proto-Bantu were affricated in Mvumbo, to /d³ tˢ/ before /i/ or to /bᵛ pᶠ/ before /u/. But plosives before non-high vocoids in Proto-Bantu were not affricated in Mvumbo, as in (14b) (from Ohala 1983, after Guthrie 1967–71).

(14)   *Proto-Bantu*   *Mvumbo*
   a.   *-buma   bᵛumo   'fruit'
        *-dib-   d³ɪwo   'shut'
        *-dut   -bᵛure   'pull'
        *-tiitʊ   tˢir   'animal'
        *-tud-   -pᶠule   'forge'
        *-gida   ma-tˢie   'blood'
        *-gubʊ   m-bᵛuu   'hippopotamus'
        *-kiŋgo   tˢiuŋ   'neck, nape'
        *-kuba   pᶠuwo   'chicken'
   b.   *-bod   -buo   'become rotten'
        *-dɪ   -di   'eat'
        *-toog   -tuog   'boil up'
        *-gada   -kala   'mat'
        *-konde   -kwande   'banana'

In contrast, the underlying plosives /t d/ in Quebec French are usually affricated into [tˢ dᶻ] only before high front vocoids. As shown in (15), the consonants /t d/ are affricated before the high front vowel /i/, the high front rounded vowel /y/, the palatal glide /j/ or the high front rounded glide /ɥ/ within a morpheme (Charbonneau & Jacques 1972, Cedergren *et al.* 1991, Ostiguy & Tousignant 1993, Papen 1998).

(15)   *Standard French*   *Quebec French*
   pe[ti]t   pe[tˢi]t   'little'
   [ti]pe   [tˢi]pe   'type'
   [tj]ens   [tˢj]ens   '(I) hold'
   [ty]rc   [tˢy]rc   'Turk'
   [tɥ]er   [tˢɥ]er   'to kill'
   [di]x   [dᶻi]x   'ten'
   [di:]re   [dᶻi:]re   'to say'
   [dj]eu   [dᶻj]eu   'God'
   [dy:]rer   [dᶻy:]rer   'to continue'
   [ty] viens le matin [tˢy] viens le matin 'you come in the morning'
   il est plain[ti]f   il est plain[tˢi]f   'he is plaintive'

Figure 34. Consonant modifier letters: f, esh, z, ezh (Kim 2001, p. 91).

Alveolo-palatal affricates in Kurdish (Suleimaniya accent)
[tɕᵘe] 'where'   [ʝzᵘe] 'ear'

Figure 35. Consonant modifier letter: turned h (Laver 1994, p. 365).

Syllabic voiceless oral stops (plosive and ejective) in Nez Perce
[tʰà + qʰɑ + k'ál +kʰ + t̩ʰ] 'to close a door'
[qʰo + qʰo + qʰeɪʔ + kʰ + t̩ʰ] 'galloped'

Figure 36. Consonant modifier letter: small glottal stop (Laver 1994, p. 240).

Laryngealization in Bwe Karen

High level tone	Mid level tone	Low level tone
[ˀwi] 'prophet'	[ˀwi] 'tasty'	–
[wa] 'finished'	[wɛ] 'rain'	[wɛ] 'smelt'

Figure 37. Consonant modifier letter: small glottal stop (Laver 1994, p. 333).

Now, if you add voicing, you can pronounce "ᵍbɑ ᵍbɑ ᵍbɑ." Do you remember the Igbo people of Western Nigeria (Biafra)? Their tribal name was usually spelled Ibo in American newspapers since outsiders seldom correctly pronounce the double consonant. You can pronounce it correctly if you say Iᵍbo — be careful not to just say Ibo, or Ig-bo.

Figure 38. Consonant modifier letter: script g (Brewster and Brewster 1976, p. 275).

Labiodentalization, which can be marked with a superscript [ᵥ], is quite common as an extralinguistic idiosyncrasy of particular individuals. In English, it is sometimes heard as a segmental feature modifying [s] and [z], and is not uncommon as a modification of [ɹ].

Figure 39. Consonant modifier letter: v-hook (Laver 1994, p. 323).

Most languages of the Iwaidjan family have a series of complex segments that have been described as "lateral flaps" (Pym and Larrimore 1979) or "prelateralized stops" (Handelsmann 1991). In all four languages apico-alveolar and apico-postalveolar complex segments /ˡt/ and /ᶩt/ exist; fuller investigation of these languages may reveal palatal /ᶩt/ as well. The complex segments contrast with simple laterals /l/ and /ɭ/, and with true clusters /lt/, /ɭt/ which span two syllables. Prelateralized stops pattern phonotactically like single phonemes. Unlike clear clusters, they can be syllable- and word-initial, as in Amurdak /ˡtaɲ/ "dingo" and /a.ˡta.wuʈ/ "water", and in slow syllabifications

Figure 40. Consonant modifier letters: l-retroflex hook, l-palatal hook (Evans 1995, p. 735).

Note that there is a typographic anomaly in the sample shown in Figure 40: retroflex (right-turning) hooks have been used on the t and modifier l, but the author was clearly discussing palatalization. What the author was intending, then, was a modifier l-palatal hook. It is not clear whether this was merely a typographic error or an attempt to approximate the palatal hook to compensate for an incomplete selection of type; it is clear, though, that the appropriate character to encode in this case is modifier l-palatal hook.

or altogether overlooked. Note also that as a consequence of rule P 6c, ǀtǀ in the first example becomes plain before ǀcǀ.
ǀpad=č'ašk+ojǀ [paᵗ'č,'aškəj] "under the cup" and ǀpod=š'ašk+ojǀ [paᵗš'aškəj]

Figure 41. Consonant modifier letter: t-palatal hook (Halle 1971, p. 71).

(21) čᵚq̣⁴² 'egg'   cʷq³ 'his mouth'
     nᵚị̃⁴² 'father'   ntʲq³⁴ 'house'

Figure 42. Consonant modifier letters: turned-m with long leg (Golston and Kehrein 1998, p. 323).

G.   References

Bailey, Charles-James N. 1985. *English phonetic transcription.* (*Summer Institute of Linguistics Publications in Linguistics*, 74.) Dallas: Summer Institute of Linguistics and University of Texas at Arlington.

Bessell, Nicola J. 1998. "Local and non-local consonant-vowel interaction in Interior Salish." Phonology 15.1–40.

Brewster, E. Thomas, and Elizabeth S. Brewster. 1976. *Language acquisition made practical: Field methods for language learners.* Colorado Springs, CO: Lingua House.

Brink, Lars; Jørn Lund, Steffen Heger, J. Normann Jørgensen. 1991. *Den Store Danske Udtaleordbog.* [Copenhagen:] Munksgaard Ordbøger.

Burenhult, Niclas. 2001. "Jahai phonology: A preliminary survey." Mon-Khmer Studies 31.29–45.

Burquest, Donald A. 2001. *Phonological analysis: a functional approach. 2nd edn., revised.* Dallas: SIL International.

Clark, John, and Colin Yallop. 1995. *An introduction to phonetics and phonology, 2nd edn.* (*Blackwell textbooks in linguistics*.) Oxford: Blackwell.

Constable, Peter. 2004a. *Revised proposal to encode phonetic symbols with palatal hook in the UCS.* L2/04-045.

——. 2004b. *Revised proposal to encode additional phonetic symbols in the UCS.* L2/04-047.

Cruttenden, Alan, ed. 2001. *Gimson's pronunciation of English. 6th edn.* London: Arnold; and New York: Oxford University Press.

Czaykowska-Higgins, Ewa, and Marie Louise Willett. 1997. "Simple syllables in Nxaʔamxcín." IJAL 63.385–411.

Evans, Nick. 1995. "Current issues in the phonology of Australian languages." *The handbook of phonological theory* (*Blackwell handbooks in linguistics*, 1), ed. by John A. Goldsmith, 723–61. Oxford: Blackwell Publishers.

Floyd, Rick. 1981. *Manual for articulatory phonetics.* Dallas: Summer Institute of Linguistics.

Gerdts, Donna. 1998. "Mapping Halkomelem voice." *Salish languages and linguistics: theoretical and descriptive perspectives* (*Trends in linguistics: studies and monographs*, 107), ed. by E. Czaykowska-Higgins and M.D. Kinkade, 305–24.

Golston, Chris, and Wolfgang Kehrein. 1998. "Mazatec onsets and nuclei." IJAL 64.311–337.

Halle, Morris. 1971. *The sound pattern of Russian.* The Hague: Mouton & Co.

Hukari, Thomas E.; Ruby Peter; and Ellen White. "Halkomelem." *Northwest coast texts* (*International journal of American linguistics native American text series, volume 2, number 3*), ed. By Barry F. Carlson, 33–68. Chicago: University of Chicago Press.

International Phonetics Association. 1999. *Handbook of the International Phonetic Association: a guide to the use of the International Phonetic Alphabet.* Cambridge: Cambridge University Press.

Kim, Hyunsoon. 2001. "A phonetically based account of phonological stop assibilation." Phonology 18.81–108.

Kraehenmann, Astrid. 2001. "Swiss German stops; geminates all over the word." Phonology 18.109–45.

Ladefoged, Peter, and Ian Maddieson. 1996. *The sounds of the world's languages.* Oxford: Blackwell Publishers.

Laver, John. 1994. *Principles of phonetics.* (*Cambridge textbooks in linguistics*.) Cambridge: Cambridge University Press.

Lojenga, Constance Kutsch. 1994. *Ngiti: a Central-Sudanic language of Africa.* (*Nilo-Saharan linguistic analyses and documentation*, 9.) Köln: Rüdiger Köppe Verlag.

Malone, Joseph L. 1999. "Some observations on the Ojibwa preverb bi." IJAL 65.343–370.

Pigott, G.L. 1997. "Licensing and alignment: a conspiracy in harmony." Phonology 14.437–77.

Pullum, Geoffrey K., and William A. Ladusaw. 1996. *Phonetic symbol guide, 2nd edn.* Chicago: University of Chicago Press.

Wayland, Ratree P., and Allard Jongman. 2001. "Chanthaburi Khmer vowels: Phonetic and phonemic analyses." Mon-Khmer Studies 31.65–82.