

White Hmong Loanword Phonology
Chris Golston & Phong Yang
California State University Fresno

Rule-based analyses of loanword phonology exhibit a peculiar trait.¹ In many instances the rules that must be posited for loans are neither rules of the donor language nor of the host language (Silverman 1992, Yip 1993). Take the Hawaiian loan *mele kalikimaka* ‘Merry Christmas’.

(1) *mele kalikimaka* ‘Merry Christmas’

m ε ɹ i k ɹ i s m ə s	English
m e l e k a l i k i m a k a	Hawaiian

A rule-based account might posit rules like the following:

(2) Rules for Hawaiian loanwords

Tensing	/lax/	->	[tense]	
Lateralization	/ɹ/	->	[l]	
Stopping	/s/	->	[k]	
Epenthesis	∅	->	V _i	/ V _i C__ {C, #}

Tensing would be responsible for changing English [ε, ɹ, ə] to Hawaiian [e, i, a]; Lateralization for replacing English [ɹ] with Hawaiian [l]; Stopping for the /s/ to [k] change; and Epenthesis for the break up of English clusters into CV sequences.

The obvious dilemma is that Tensing, Lateralization, Stopping and Epenthesis are not and could not be rules of Hawaiian because Hawaiian has no lax vowels, ɹ, s or consonant clusters for such rules to apply to. But these are clearly not rules of English either, for English leaves lax vowels, ɹ, s and consonant clusters unchanged. Thus the loanword phonology requires rules that are neither rules of the donor language (English) nor of the native language (Hawaiian). We could argue that loanword phonology uses Rules of Universal Grammar, but that would be a short-lived solution as well: different languages adapt loanwords differently, not according to one grand

¹ We would like to thank our Hmong consultants for their linguistic expertise: Chue Her, Neng Thao, Paly Yang, Lou Xiong.

scheme. *Christmas* is borrowed into White Hmong as *k^hĩmǎ* not as *kalikimaka*.

Rule-based analyses like this miss the generalization that when a language adopts a loan-word into its vocabulary it attempts to bring that word into conformity with the phonology of the language (Yip 1993). The facts here are similar to those of second language acquisition, where hypothesized rules of the interlanguage are often rules neither of the native language nor of the target language (Broselow, Chen & Wang 1998; Alber & Plag to appear). Constraint-based approaches to are better suited to explaining both phenomena, as Yip has shown for loanwords and Broselow, Chen & Wang have shown for second language acquisition. In both cases, the constraints that comprise the native phonology can be used to understand how loanwords are nativized and to characterize the negative transfer we identify as a foreign accent.

In this paper we will analyze the loanword phonology of White Hmong in terms of Optimality Theory (Prince & Smolensky 1993, McCarthy & Prince 1993ab), following Yip's work on Mandarin loanword phonology. Our data comes from two sources: French loans, including a large number of biblical terms from a Hmong bible printed in France, and English loans, collected from a number of speakers in central California where a large Hmong immigrant population has resided since the 1980s. The analysis should translate straightforwardly into other constraint-based formalisms but we use OT as a *lingua franca*. Our main interest is in what can be modeled well by constraint-based formalisms *in general* as opposed to rule-based formalisms *in general*.

After briefly reviewing the phonology of White Hmong (§1), we turn to loanword data concerning sounds (§2) and syllables (§3), where our results basically parallel those of Yip (1993). This all works out quite well. Where our constraint-based analysis does not fare well is in tones (§4) and in polysyllabic words (§5). Here we find that toneless words from English and French show up with only 4 of the 7 tones we find in Hmong (low, falling, rising, creaky): no loanwords have the H or M tones so common in Hmong and none has breathy voice. This cannot be attributed to the constraints of Hmong, since they freely permit H, M, and breathy. We have no solution to this problem yet and leave it to future research.

The second area we encounter problems is in polysyllabic words. Hmong has none of these (except for compounds), so our constraint-based approach predicts that loanwords will shorten to monosyllables. But they do not. Polysyllables are freely loaned into Hmong as polysyllables. We

propose that such words are borrowed in as compounds. This is descriptively adequate but not yet fully satisfying, as we have not yet found any corroborating evidence (e.g., in tone-sandhi rules) that supports our hypothesis. So this remains a problem as well for future research.

We conclude that a constraint-based analysis is decidedly superior to any conceivable rule-based analysis (which must use rules that are neither part of Hmong nor part of the donor language) but not yet fully adequate. Some structures that are impossible in Hmong (polysyllabic morphemes) are freely allowed in loans, and some structures that are freely allowed in Hmong (H, M, breathy tones) are ruled out in loanwords. As far as we know this will be a problem for *any* theory of loanword phonology that uses the patterns of the native language to model the changes made to loans.

1. A sketch of White Hmong phonology

Hmong is a Miao-Yiao language traditionally spoken in southern China, Thailand and Laos. It has two main dialects: White Hmong and Green² Mong (without the initial H). The language is consistently monosyllabic and isolating: every morpheme is one syllable long and every word has a one morpheme (modulo compounds). Thus our discussion of Hmong syllables below covers the shapes of Hmong morphemes exactly. Hmong has no codas to speak of, but very ornate onsets and nuclei, with a rich system of tones and laryngeal contrasts. We concentrate here on White Hmong, which we will simply call Hmong, reserving Mong for the Green dialect of the language.

1.1 Sounds

Hmong has a large number of consonants including some 39 stops. These occur at six places of articulation (labial, dental, alveolar, palatal, velar, uvular). Some analyses of Hmong treat some of these sounds as clusters ([m̥] as [hm], [p^h] as [phl], etc.) but we will treat them as simple sounds following Golston & Yang 2000. For present purposes the effects of this decision are nil.

² For recent discussion of phonation types and tone in Green Mong see Andruski & Ratliff 2000.

(3) Consonants (plus ʔ and h)

	Lab	Dent	Retr	Pal	Vel	Uv
Stop	p p ^h	t t ^h d d ^h	ʈ ʈ ^h	c c ^h	k k ^h	q q ^h
lateral	p ^l p ^ʎ					
prenas	^m b ^m b ^h	ⁿ d ⁿ d ^h	ⁿ ʈ ⁿ ʈ ^h	ⁿ ʃ ⁿ ʃ ^h	^ŋ g ^ŋ g ^h	^N G ^N G ^h
strident		t ^s t ^{sh}	tʃ tʃ ^h			
prenas-	^m b ^l ^m b ^ʎ					
lateral						
prenas-strid		ⁿ d ^s ⁿ d ^{sh}	ⁿ dʒ ⁿ dʒ ^h			
Fricative	v f	s	ʒ ʃ	ʝ ʥ		
Nasal	m m̥	n n̥		ɲ ɲ̥		
lateral	m ^l m̥ ^ʎ					
Approx		l ʎ				

Following Kehrein (2000), we assume that phonetic affricates are stops phonologically and that stops may be either simple, [strident], [lateral] or [nasal] or combinations of these. Hmong exhibits a 6-way contrast, distinguishing simple (p, t), lateral (p^l), strident (t^s), nasal (^mp), nasal-lateral (^mp^l), and nasal-strident (ⁿt^s), each in a plain (p, p^l, t^s, ^mb, ^mb^l, ⁿt^s) and an aspirated series (p^h, p^ʎ, t^{sh}, ^mb^h, ^mb^ʎ, ⁿd^{sh}). Note that lateral release only occurs with labials and that strident release only occurs with coronals:

(4) Stop contrasts

	Labial		Coronal
simple	pua ‘cover’		tɔ ‘bleed’
lateral	p ^l ua CLASSIFIER		
strident			t ^s ɔ ‘residue’
nasal	^m bua ‘pig’		ⁿ dɔ ‘reach the top’
nasal-lateral	^m b ^l ua ‘slippery’		
nasal-strident			ⁿ d ^s ɔ ‘bite’

A slightly fuller system is found in Mong, where the equivalent of Hmong voiced [d] is a laterally released [t^l]. (The use of contrastive voicing in

Hmong is limited to *d*, *d^h*.) We note in passing that prenasalized stops show up voiced during stop closure. This will become important when we see how voiced stops are borrowed into Hmong.

Hmong vowels are somewhat less daunting, at least if we abstract away from the rich system of tones, which we'll treat separately below. Hmong has three high vowels, a low vowel and four mid vowels, two of which are nasal:

(5) Vowels

	Front	Central	Back
Hig h	i	ɨ	u
Mid	e ẽ		ɔ õ
Low		a	

Hmong also has five diphthongs, three falling in sonority, two rising:

(6) Diphthongs

Falling	Rising
ai	ia
ai	
au	ua

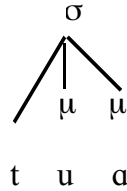
The rising diphthongs are somewhat centering (*iə*, *uə*), but we treat them here as [*ia*] and [*ua*] since Hmong has no mid central vowels otherwise.

1.2 Syllables

Given the rich array of consonants just discussed, the syllable structure of Hmong is extremely simple, consisting of a mandatory onset (in all but a few particles) and one or two vowels CV(V). With the exception of glottalized nuclei, which are cut fairly short, all Hmong syllables are roughly isochronous, there being no distinction between the length of a monophthong and a diphthong. Rising and falling tones can occur on both monophthongs and diphthongs. For this reasons we have chosen to represent all Hmong syllables as bimoraic.

The Hmong syllable, then, is well represented by the moraicly based model found in McCarthy & Prince 1986, Hayes 1989 and elsewhere:

(7) Hmong syllables (and thus Hmong morphemes)



Again, it must be stressed that Hmong onsets contain no real clusters and that Hmong syllables have no codas. Two sounds, [ŋ] and [ʔ], may occur in the coda, but in each case they are in free-variation: ‘phonologically there is no final consonant, but nasalization...is often realized as final [ŋ] and tone /m/ (i.e., creaky voice, CG & PY) often has an accompanying non-distinctive final [ʔ]’ (Heimbach 1966, xiii). To keep diacritics somewhat at bay we will transcribe nasalized vowels as Vŋ, but we do not mean to imply that the nasal stop is in any way required. On the contrary, nasalization on the vowel is quite stable while the final nasal stop comes and goes, just as creaky voice is stable but the final glottal stop comes and goes.

1.3 Tones

Traditional analysis of Hmong tones includes seven basic ‘tones’, which Yang (2000) analyses as follows:

(8) Hmong tones

p ^ó	‘lump’	t ^ó	‘deep’
p ^ǝ	‘pancreas’	t ^ǝ	‘bleed’
p ^ḁ	‘thorn’	t ^ḁ	‘wait’
p ^ḡ	‘female’	t ^ḡ	‘hill’
p ^ǎ	‘throw’	t ^ǎ	‘mix’
p ^ḙ	‘see’	t ^ḙ	‘bite’
p ^{ḙḙ}	‘grandmother’	t ^{ḙḙ}	‘sink’

Distributional considerations lead us to think that M is the unmarked tone, but we will see that the loanword phonology has some bearing on this. The last two ‘tones’ are analyzed here as breathy and creaky voice, their primary acoustic cues (Yang 2000).

This brings our little overview of Hmong phonology to a close. It is worth pointing out that Hmong has nothing that one might call a phonological rule ($x \rightarrow y / a_b$) except a minor tone sandhi rule that will not concern us here because its structural description is never met in Hmong loans. Except for some fast speech phenomena, Hmong morphemes surface with all the sounds and tones that they have underlyingly. With this in mind we can now turn to Hmong loanword phonology, which seems to be alive with rules deleting consonants, epenthesising vowels and changing features.

2. Loanword Sounds

Not surprisingly, the sounds we find in Hmong loanwords are exactly those we find in native words. Borrowing an idea from Aronoff (1976, 98) and Kiparsky (1982) we call this structure preservation:

(9) STRPRES Only contrastive sounds of the language are allowed.

When all the sounds of the borrowed item are found in Hmong, and when the borrowed syllables are of the shape CV, we get exact borrowings, modulo tone.

(10) Exact borrowings (modulo tone)

dà.mà	< da.ma	‘Damascus’	(<French)
dè.kà.pò.ḽi	< de.ka.po.li	‘Decapolis’	
fi.lè.mòḽ	< fi.le.mõ	‘Philemon’	
mǐ.kà	< mi.ka	‘Micah’	
sǐ.nǎi	< si.nai	‘Sinai’	
tǐ.mò.tè	< ti.mo.te	‘Timothy’	
t ^h i.vǐ	< t ^h i.vi	‘T.V.’	(<English)
ʃḽ	< ʃḽ	‘Shaw’	
mà.nà	< ma.nə	‘manna’	
mǔ.vǐ	< mu.vi	‘movie’	

Following work in OT, we attribute the exactness of the borrowings to the faithfulness constraints MAX and DEP (McCarthy & Prince 1995).

(11) MAX No deletion.

(12) DEP No insertion.

But most loans are not so lucky, and when the source word contains sounds not found in Hmong the loan word surfaces with sounds that are. Generally, the sounds we find in Hmong loans from English are simply the sounds that English and Hmong share; the sounds that we find in Hmong loans from French are simply the sounds that French and Hmong share. This is a straightforward effect of structure preservation. Things are not quite so simple, though, as sometimes Hmong loans have sounds that English and French lack. In these cases, the Hmong sound is the closest equivalent to the English or French sound in terms of features, as we will see.

Hmong has no voiced consonants other than [d] and [d^h], so voiced labials and velars in French and English are borrowed as prenasalized in Hmong loans. We see this, for instance, with voiced labial and velar stops, both initial and medial:

(13) ^mb < b

^m bà.là. ^m bà	ba.ɰa.bas	‘Barabas’	(<Fr)
^m bà.lɿ		‘Baruch’	
^m bà. ^m bè	ba.bɛl	‘Babel’	
^m bà. ^m bĩ.lò.	ba.bi.lɔn	‘Babilonia’	
^m bè.lè.hè	bɛt.le.ɛm	‘Bethlehem’	
^m bè.sàì.dà		‘Bethsaida’	
^m bè.tà.nĩ	be.ta.ni	‘Bethany’	
hà. ^m bà.kĩ	a.ba.kyk	‘Habakkuk’	
à.là. ^m bĩa	a.ɰa.bi	‘Arabia’	
à. ^m bè	a.bɛl	‘Abel’	
è.lĩ.sà. ^m bè	e.li.za.bɛt	‘Elizabeth’	
hà. ^m b ^l à	a.bɰa.am	‘Abraham’	
hè. ^m b ^l ǎ	e.bɰø	‘Hebrew’	
hè. ^m b ^l ɔŋ	e.bɰõ	‘Hebron’	

li. ^m bà	li.bã	‘Lebanon’	
mà.kà. ^m be	ma.ka.be	‘Maccabees’	
^m bïà	bɪ	‘beer’	(<Eng)
^m bǎ	bəs	‘bus’	
^m bêŋ	bæŋk	‘bank’	
^m bǒ	baks	‘box’	
^m bǒ.daj̄	bə.rəɪ	‘butter’	
^m bǎi. ^m bɔ	bai.bəl	‘bible’	
hà. ^m bǎi. ^ŋ gǎi	hæm.bəɪ.gəɪ	‘hamburger’	

(14) ^ŋg < g

^ŋ gà.là.ɸià	ga.la.tian	‘Galatians’	(<Fr)
^ŋ gà.li.lài	ga.li.le	‘Galilee’	
^ŋ gà. ^m b ^k ia	ga.bɤjɛl	‘Gabriel’	
ŋgà.ntsà	ga.za	‘Gaza’	
à.ŋgǎi	a.ge	‘Haggai’	
hài.ŋgǐ.tò	e.ʒipt	‘Egypt’	
ŋgě	gæs	‘gas’	(<Eng)

We assume that prenasalization occurs in order to preserve the strong voicing found in French or the weaker voicing found in English:

(15) ^ŋgà.ⁿd^za < ga.za (Fr.)

/ga.za/	STRPRES	MAXVOI	DEPNAS
gà.zà	*!*		
kà.sa		*!*	
^ŋ gà. ⁿ d ^z a			**

The most faithful candidate [gà.zà] loses because Hmong has no voiced [g] in its inventory. This is avoided in the second candidate [kà.sà] but at a cost: the voicing in the source word is lost, in violation of MaxVoi. The final candidate violates lower-ranked DEP_{NAS}, but this is the best solution given the grammar of the language.

We do not find prenasalization of this type with sourceword which contain [d]. White Hmong has a voiced [d], so that sound is borrowed as is rather than as a prenasalized [ʰd], which would violate Dep_{Nas} gratuitously:

(16) d < d

dà.nǎ	da.nɛl	‘Daniel’	(<Fr)
dà.vǐ	da.vid	‘David’	
mà.se.dò.nǎ	ma.se.dwan	‘Macedonia’	
jù.dà	ʒud	‘Jude’	
si.dõŋ	si.dõ	‘Sidon’	
jò.la.dà	ʒuɤ.dẽ	‘Jordan’	
à.dà	a.dam	‘Adam’	
mè.dǐ.tè.là.nè	me.di.te.ɤa.ne	‘Mediterranean’	
dò.là	da.ləɪ	‘dollar’	(<Eng)
sǐ.dǎi	si.dəɪ	‘Cedar’	

Again, the reason for the retention of [d] is that Hmong has [d] onsets already. Maintaining the voicing in the loan does not violate structure preservation, so the first candidate wins: the competition either loses the voicing or adds nasalization for no reason:

(17) dà.vǐ < da.vid (Fr.)

/da.vid/	STRPRES	MAXVOI	DEPNAS
dà.vǐ			
tà.fǐ		*!*	
ⁿ dà.vǐ			*!*

We deal with the loss of the word-final [d] below and ignore it for the time being.

Another clear case of structure preservation is the replacement of French [ʁ] and English [ɹ] by Hmong [ɿ]:

(18) ɿ < ʁ, ɹ

lǔ	ɿyt	‘Ruth’	(<Fr)
lò.mà	ɿo.mě	‘Romans’	
fà.là.ɲ	fa.ɿa.ɲ	‘Pharaoh’	
fà.ɿi.sàì	fa.ɿi.zjě	‘Pharisee’	
sà.là	sa. ɿa	‘Sarah’	
sà.mà.ɿià	sa.ma.ɿi	‘Samaria’	
sè. ⁿ d ^s à.lè	se.za.ɿe	‘Caesarea’	
hǐ.sà.lǎì	iz.ɿa.el	‘Israel’	
nà.sà.lè	na.za.ɿet	‘Nazareth’	
à.mè.ɿi.kà	ə.mɛɿ.i.kn	‘American’	(<Eng)
t ^h ù.là.ɿi	t ^h u.lɛɿ.i	‘Tulare’	

Our general point is that Structure Preservation does most of the segmental work in our analysis. That is, we explain the sounds used in Hmong loanwords by means of a set of constraints already found in Hmong. A rule-based analysis must posit a number of superfluous rules for Hmong loanword phonology that are not found in native Hmong, or in French, or in English. And this for a language that has no phonological rules that convert one set of sounds into another! For the sake of completeness, we list the other cases we’ve isolated:

(19) f < f

fè.nǐ.sǎ	fe.ni.siě	‘Phoenicia’	(<Fr)
fi.ɿi.pò	fi.li.pjě	‘Philippians’	

³ Presumably from ‘Roma’, hence the final [a].

fɔŋ	fon	‘phone’	(<Eng)
ʔə.fi	ɔ.fis	‘office’	
tè.lè.fɔŋ	t ^h ɛ.lə.fon	‘telephone’	

(20) v < v

ʔè.và	ɛ.va	‘Eve’	(<Fr)
mǔ.vj	mu.vi	‘movie’	(<Eng)

(21) s < s

sà.lò.mòŋ	sa.lo.mõ	‘Salomon’	(<Fr)
sà.lè.tà		‘Zarephath’	
sà.mùà	sa.myel	‘Samuel’	
sà.tà	sa.tã	‘Satan’	
sò.fò.nià	so.fo.ni	‘Zephaniah’	
sè. ⁿ d ^s à	se.za	‘Caesar’	
sì.nài	si.nai	‘Sinai’	
si.dòŋ	si.dõ	‘Sidon’	
hì.sài	i.sai	‘Isaiah’	
kò.lò.sài	ko.lo.siẽ	‘Colossians’	
sài.kû	skul	‘school’	(<Eng)
sì.dãi	si.dɿ	‘Cedar’	
sǒ.fạ	so.fə	‘sofa’	
kà.sě	k ^h ə.set	‘cassette’	
pě.sî	pɛp.si	‘Pepsi’	

(22) s < z (no z in Hmong)

ʔè.sè.kià	e.ze.kjel	‘Ezekiel’	(<Fr)
ʔè.fè.sò	e.fe.zjẽ	‘Ephesians’	
hè.sà.hì	e.zaj	‘Esau’	

hĩ.sà.hà	i.zak	‘Isaac’
jò.sè	ʒo.zɛf	‘Joseph’
sè. ⁿ d ^s à.lè	se.za.ɕe	‘Caesarea’
sè. ⁿ d ^s à	se.za	‘Caesar’
ⁿ d ^s à.kà.fìa	za.ka.ɕi	‘Zechariah’

(We have no explanation for why some [z] words are loaned as [s] while others are loaned as [ⁿd^s]. Perhaps more data will shed some light on this in the future.)

(23) ʃ < ʒ

ʃò	ʒo	‘Shaw’	(<Eng)
mĩ.ʃi. ^m bì.ʃi	mit.sə.bi.ʃi	‘Mitsubishi’	

(No explanation for ʃ < t.s)

(24) j < ʒ

jà.kò	ʒak	‘James’	(<Fr)
jà.kò	ʒa.kɔb	‘Jacob’	
jò.è	ʒo.ɛl	‘Joel’	
jò.nà	ʒo.nas	‘Jonah’	
jò.sè	ʒo.zɛf	‘Joseph’	
jò.ʃùà	ʒo.sye	‘Joshua’	
jè.lè.mĩa	je.ɕe.mi	‘Jeremiah’	

(Our analysis incorrectly predicts that French ʒ would be borrowed as such, since Hmong has that sound. But this is not the case, something we have no explanation for at present.)

(25) tʃ^h < tʃ

tʃ ^h âi	tʃə.tʃ	‘church’	(<Eng)
tʃ ^h î	tʃiz	‘cheese’	
ⁿ dʃû	dʒus	‘juice’	

Some sounds of Hmong never show up in loanwords, but we presume this is because there are no sounds in French or English similar enough to call them forth. The full list of consonants is:

(26) Consonants not found in Hmong loanwords

q, q^h, ⁿq, ⁿq^h
 ml, mɬ
 m̩, n̩, ɲ, ɳ
 ɬ
 j, ʒ

This can be explained by the fact that neither French or English have these sounds to start with and need not concern us further.

Turning now to vowels, Hmong loanwords are again quite straightforward: lax vowels that Hmong doesn’t have are borrowed in as the vowel in Hmong that is closest to them in terms of height, rounding and backness.

(27) a < a, æ, ə

kàfê	kafe	‘coffee’	(<Fr)
jà.kò	ʒak	‘James’	
jà.kò	ʒa.kɔb	‘Jacob’	
ʔà.p ^h ǎ.mèŋ	ə.p ^h a.ɪ.t.mɪnt	‘apartment’	(<Eng)
ʔò.tò.mà.fǐ	ɔ.to.mæ.tɪk	‘automatic’	
tʃ ^h à.nò	tʃ ^h æ.nɪ	‘channel’	
ʔà.mè.fi.kà	ə.mɛɪ.ɪ.kə	‘America’	

(28) e < e, ε

kàfê	kafɛ	‘coffee’	(<Fr)
jè.lè.mìà	je.ɤe.mi	‘Jeremiah’	
jè.lì.kò	ʒe.ɤi.ko	‘Jerico’	
mbà.ŋkě	ba.ɣɛt	‘baguette’	
mè.t ^h ò	mɛ.tɤo	‘metro’	
ʔè.t ^h è	ɛs.tɛɤ	‘Esther’	
jò.sè	ʒo.zɛf	‘Joseph’	
k ^h ê	k ^h ɛk	‘cake’	(<Eng)
pě.sî	pɛp.si	‘Pepsi’	

(29) eŋ < ẽ

pèŋ	pẽ	‘bread’	(<Fr)
kà.èŋ	ka.ẽ	‘Cain’	

(There is of course no change in nasalization here, since our <eŋ> represents a nasalized vowel; the change comes in the vowel height: low-mid for French, hi-mid for Hmong).

(30) i < i, ɪ

ṭè.sà.lò.nǐ	te.sa.lo.ni.siẽ	‘Thessalonians’	(<Fr)
jè.lì.kò	ʒe.ɤi.ko	‘Jerico’	
mě.k ^h ɪŋ.lǐ	mək.kɪ.li	‘McKinley’	(<Eng)
tʃ ^h ɪ	tʃiz	‘cheese’	
mǐ.ʃi. ^m bǐ.ʃǐ	mɪt.sə.bi.ʃi	‘Mitsubishi’	
mě.k ^h ɪŋ.lǐ	mək.kɪ.li	‘McKinley’	
kɫò.vǐ	kɫo.vɪs	‘Clovis’	

(31) u < u, y

lù.kà	lyk	‘Luke’	(<Fr)
jè.lù.sà.lè	ʒe.ɤy.za.lɛm	‘Jerusalem’	
jè.sù	ʒe.zy	‘Jesus’	
ⁿ tʒû	dʒus	‘juice’	(<Eng)
t ^h ù.là.lì	t ^h u.lɛ.i	‘Tulare’	

(32) ɔ < ɔ, o

ʔð.fì	ɔ.fis	‘office’	(<Eng)
ʔð.lí	ɔ.liv	‘Olive’	
dò.là	dɔ.lər	‘dollar’	
ʔð.tù.mà.tǐ	ɔ.to.mæ.tɪk	‘automatic’	
ʔð.jò.tà	t ^h o.jo.ta	‘Toyota’	

(33) Diphthongs

p ^h ô	p ^h aɪk	‘park’	(<Eng)
tʃ ^h âi	tʃəɪtʃ	‘church’	
háù	haus	‘house’	
âi	ais	‘ice’	
hòŋ.tà	han.də	‘Honda’	

Before we leave loanword vowels, we note that there is one vowel, [i], that is found in Hmong but not in our corpus of loanwords (except as part of the diphthong ai.). We assume that no sound of English or French is similar enough to call this sound forth in Hmong.

Generally, we find that structure preservation does a good job of accounting for differences between Hmong loans and the source words they come from. Since there are no native phonological rules that could convert non-native sounds to native sounds, a rule-based analysis of such data is extremely ad hoc, unless of course, it is driven by constraints. But rules alone cannot account for such data in a satisfying manner.

3. Loanword Syllables

One of the most obvious accommodations loanwords into Hmong make is in terms of syllable structure. Whatever the input syllables, the output syllables are CV(V). Vowel-initial words, for instance, pick up an initial glottal stop when they are borrowed into Hmong from English or (for the most part) French:

(34) ? inserted with vowel-initial words

ʔà.lòŋ	a.Rɔ̃	‘Aaron’	(<Fr)
ʔà.lè.sà. ⁿ cǎ	a.lek.sã.dɾi	‘Alexandria’	
ʔà.là. ^m bǎ	aRabi	‘Arabia’	
ʔà.lòŋ	a.Rɔ̃	‘Aaron’	
ʔà. ^ŋ kǎi	a.ge	‘Haggai’	
ʔè. ⁿ cà	es.dɾa	‘Ezra’	
ʔè.fè.sò	e.fe.zjẽ	‘Ephesians’	
ʔè.sè.kǎ	e.ze.kjɛl	‘Ezekiel’	
ʔè.tè	es.tɛR	‘Esther’	
ʔè.và	ɛ.vɑ	‘Eve’	
ʔi.tà.lǎ	i.ta.li	‘Italie’	
hè.lò	e.Rɔd	‘Herrod’	?
hè.mòŋ		‘Hermon’	?
hǎi.ŋkǎ.tò	e.zipt	‘Egypt’	?
ʔà.p ^h ǎ.mèn	ə.p ^h a.t.mɪnt	‘apartment’	(<Eng)
ʔà.mè.li.kà	ə.mɛɪ.ɪ.kə	‘America’	
ʔò.fǐ	ɔ.fɪs	‘office’	
ʔò.tò.mà.tǐ	ɔ.to.mæ.tɪk	‘automatic’	
ʔǎi	ais	‘ice’	

(We have no explanation at present for loans from French that surface with initial [h], but assume it has something to do with crazy French spelling.) We attribute the epenthesis of glottal stop to the presence (in Hmong and

probably all languages) of a constraint requiring syllables to begin with consonants:

(35) ONSET: Syllables begin with consonants. (Prince & Smolensky 1993)

More interesting, perhaps, is the fact that syllable-final consonants are categorically deleted:

(36) deletion of syllable-final consonants

lǔ	ryt	‘Ruth’	(<Fr)
jó	ʒob	‘Job’	
ǎi.tò	ti.tos	‘Titus’	
nǐ	nil	‘Nile’	
jè.lù.sà.lè	ʒe.ri.za.lem	‘Jerusalem’	
ʔà.dà	a.dam	‘Adam’	
c ^h ùá	tʃɹæk	‘truck’	(<Eng)
k ^h ê	kek	‘cake’	
mpǒ	baks	‘box’	
k ^h è.mâ	ke maɪt	‘KMart	
sè. mâ	sev maɪt	‘SaveMart	
wà.mâ	wal maɪt	‘WalMart	

We attribute this to a constraint of Hmong (and languages generally) banning syllabl-final consonants:

(37) NOCODA: Syllables end with vowels. (Prince & Smolensky 1993)

We see, then, that loanword syllables are well-modeled in a constraint-based analysis because there are well-founded native constraints that account for the simplification of complex syllables found in the source languages. The same data is poorly modeled with rules because Hmong lacks any such rules.

But we must not be too happy about the result. As far as we can tell, OT has no way of explaining why Hmong loans satisfy NOCODA by deleting consonants rather than by epenthesizing vowels. We might stipulate that

Dep is more highly ranked than Max, but the question will then merely be passed on to the acquisition process: how would Hmong children learn that Dep is more highly ranked than Max without any data in their native language that bears on the issue? We have toyed with the idea that the monosyllabism we find in Hmong might lead us out of this mess, since it would weigh-in on the side of deletion rather than epenthesis; but Emmanuel Dupoux rightly points out (p.c.) that this still requires the Hmong child to know something about how that constraint is ranked with respect to Max and Dep and that there is no data in Hmong itself that would help the child learn such a ranking.

4. Tones

Turning now to tones, we encounter some major problems with a constraint-based analysis of loanword phonology. The embarrassing fact is that the tones we find in Hmong loans seem to be a completely arbitrary set of the tones found in Hmong. Furthermore, we find different patterns in French and English loans that does not seem to be due to the phonology of French or English and cannot be due to the phonology of Hmong. Let us begin with this latter fact.

There is a striking asymmetry in tones between loans from French and loans from English. All of the French loans we collected have Low tones all the way across the word. Consider

(38) Loans from French surface with Low tone

kà.nà.hà	ka.naã	‘Canaan’
kò.è.lè	e.kli.zi.ast	‘Ecclesiastes’
kò.li.nt ^h ò	ko.Rēt	‘Corinthians’
mà.kò	mark	‘Mark’
mà.là.kǐà	ma.la.fi	‘Malachi’
mǎi.sè	mo.iz	‘Moses’
má.líá	ma.Ri	‘Marie’
mà.tǎi	ma.tjø	‘Matthew’
mè.sò.pò.ṭà.mǐà	me.so.po.ta.mi	‘Mesopotamia’
nò.hè	no.a	‘Noah’

nè.hè.mǐ	ne.e.mi	‘Nehemiah’
nǐ	nil	‘Nile’
pǐ.là.tò	pi.lat	‘Pilate’
tǐ.tò	ti.tos	‘Titus’
tǐ.lò	tiR	‘Tyre’
jè.fi.kò	ʒe.Ri.ko	‘Jericho’
jè.lù.sà.lè	ʒe.Ri.za.lɛm	‘Jerusalem’

That is:

(39) Hmong loanword tones from French

H	— ⁴
M	—
L	all
HL	—
LH	—
cg	—
sg	—

We can imagine all sorts of interesting stories to make of this. An account in terms of underspecification might claim that L tone is unmarked in Hmong and since French words don’t bear contrastive tone they all enter Hmong with the unmarked tone, hence with L. Or an OT account might rank the constraint that keeps low tones out of French very low in Hmong, so that Hmong speakers would rather insert it than a tone whose presence violates a higher ranked constraint. And so on. The problem here is not in coming up with some sort of analysis.

The problems is that loans from English do not all enter as L in Hmong. Rather, we find that different syllable structure in the borrowed item induces different tones in the loans. Indeed, there is a surprisingly uneven distribution of tones in English loanwords. Some English words enter Hmong with a low tone (L), some with falling (HL) or rising (LH) and a few with creaky voice (cg). But we have found no English loans with breathy voice, mid (M) or high (H) tone.

⁴ We have two exceptions in our corpus, neither of them compelling. Peter [pó.ʒé], literally ‘rock’, hence the tone. Job is [jǒ] but the H tone here may be orthographically driven (final –b marks H in Hmong spelling).

That is, instead of (39) we find the following for English loans:

(40) Hmong loanword tones from English

H	—
M	—
L	many
HL	many
LH	many
cg	few
sg	—

The tones do not seem to be randomly distributed and we find the following general patterns, marred by a few exceptions.

(41) creaky < final stressless syllable in English

sǐ.tǔ.pɛ̄	stu.pɪd	‘stupid’	
sǒ.fǎ̄	so.fə	‘sofa’	
mǔ.vᵢ	mu.vi	‘movie’	
sǐ.tě.pɔ̄	ste.pl	‘staple’	
mě.pǎᵛ	me.pəl	‘Maple’	
mě.pɔ̄	me.pl	‘Maple’	
mǎ.tǎ.nɔ̄	mək.da.nəldz	‘McDonald’s’	
^m bǎī. ^m bɔ̄	bai.bl	‘bible’	
t̃i.tʃ ^h aī	t ^h i.tʃ ^h əɪ	‘teacher’	
^m bǒ̄.tāī	bə.tɪ	‘butter’	
kòŋ.p ^h ǔ̄.tāī	kəm.pju.tɪ	‘computer’	
mě.k ^h ɔ̄	mək.k ^h al	‘McCall’	(?)

The final word here has creaky voice but does not end in a stressless syllable in English. The next pattern involves English stressed syllables with long nuclei (tense vowels, diphthongs or Vɪ):

(42) HL: VV

ʔâi	aɪs	‘ice’	VV
háù	haus	‘house’	
k ^h ê	kek	‘cake’	
ndzû	dʒus	‘juice’	
tʃ ^h î	tʃiz	‘cheese’	
k ^h è.mâ	k ^h e.maɪt	‘Kmart’	Vɪ
sè. mâ	sev.maɪt	‘Savemart’	
wà.mâ	wal.maɪt	‘Walmart’	
p ^h â	p ^h aɪk	‘park’	
p ^h â.kèŋ	paɪ.kɪŋ	‘parking’	
tʃ ^h âi	tʃəɪtʃ	‘church’	
sì.tâi	si.dɪ	‘Cedar’	?
wû.wâi	wud.wɪd	‘Woodward’	?
ʔò.lî	a.lɪv	‘Olive’	?
^m bêŋ	bæŋk	‘bank’	?

The last four words are exceptional. We’d expect the first two to surface with final creaky voice (see above); the third with a rising tone (see below); and the last with a falling tone (see below).

Rising tone is generally found on syllables that end in voiceless consonants in English. Consider the first group of words below:

(43) LH: V + vcls C

pě.sî	pɛp.si	‘pepsi’	
sì.tǒ	stap	‘stop’	
kà.sě	kə.set	‘cassette’	
ʔò.fî	ɔ.fɪs	‘office’	
kɫò.vǐ	kɫo.vɪs	‘Clovis’	
^m bǎ	bəs	‘bus’	‘
^ŋ gě	gæs	‘gas’	
p ^h è.lě	p ^h e.lɛs	‘Payless’	

pù.lǐ	pə.lis	‘police’	
k ^h à.pě	kaɪ.pət	‘carpet’	
tʃ ^h ě.sà.nǎ	tʃest.nət	‘Chestnut’	
c ^h ùá	tʃɹək	‘truck’	
p ^h ǐ.nǐ	p ^h ɪk.nɪk	‘picnic’	
sǐ.tǒ.tǎŋ	stak.tən	‘Stockton’	
tʃ ^h ě	tʃ ^h ɛk	‘check’	
^m bǒ	baks	‘box	‘
mě.k ^h ǒ	mək.k ^h al	‘McCall’	
mě.k ^h ǐ.li	mək.kɪn.li	‘McKinley’	
mǎ.tǎ.nǒ	mək.da.nəldz	‘McDonald’s’	
sǒ.fǎ	so.fə	‘sofa’	?
sǐ.tě.pǒ	ste.pl	‘staple’	?
^m bǎi. ^m bǒ	bai.bl	‘bible’	?
mǔ.vǐ	mu.vi	‘movie’	?
sǎ.lǎ	sæl.ɪd	‘salad’	?

The second group of words is somewhat problematic (8 cases). The source syllable in these words does not end in a voiceless consonant, although the next syllable does begin with one. If such words are produced with the voiceless consonant ambisyllabic (as is often claimed for American English) perhaps they are perceived in coda position by Hmong speakers. If not, they are exceptional. The last three cases are not consistent with our analysis. We’d expect the first syllable of ‘bible’ and ‘movie’ to be falling. We’d expect the first syllable of ‘salad’ to be L and the second to be creaky.

Finally, L tone is generally found where the English source has an open syllable, a nasal-final syllable, or a syllabic liquid. We take this to be the elsewhere case, as suggested by loanwords from French:

(44) L: open syllable, __nas

ʃǒ	ʃǒ	‘Shaw’	open syllable
ʔà.mè.li.kà	ə.mɛ.ɪ.kə	‘America’	

ʔò.tò.mà.tǐ	ɔ.to.mæ.tɪk	‘automatic’	
mǐ.ʃi.bǐ.ʃi	mɪt.sɪ.bi.ʃi	‘Mitsubishi’	
tʰi.vǐ	ti.vi	‘T.V.’	
tò.jò.tà	to.jo.ta	‘Toyota’	
tè.lè.fɔŋ	tɛ.lə.fon	‘telephone’	
tʰù.là.lǐ	tə.lɛ..li	‘Tulare’	
fɔŋ	fon	‘phone’	__nas
hòŋ.dà	han.də	‘Honda’	
sǐ.tèŋ	stæmp	‘stamp’	
tʃʰà.nò	tʃæ.nl	‘channel’	syllabic liquid
dò.là	dɔ.lɪ	‘dollar’	
hà. ^m bǎi. ^ŋ gǎi	hæm.bɪ.gɪ	‘hamburger’	
^m bǎi	bɪ	‘beer’	?

‘Beer’ is problematic here, unless it is perceived as two syllables by Hmong listeners [bi.ɪ]. The local dialect of English is not inconsistent with this possibility.

The details of the processes here should not detract from the major problems these patterns cause for our understanding of Hmong loanword phonology. We see two major areas of difficulty. The first concerns the difference between French and English loans. Why should the former surface as L all the time while the latter surface as creaky, rising, falling or L depending on the phonology of the English word?

More distressing is the fact that the tones in the loans co-vary with sounds in English that do not appear in the loans themselves. This raises problems for every framework we are aware of, including the one we adopt here, OT.

In a surface-oriented framework like OT, we must ask how a consonant that does not surface can affect the surface tone. Even if we come up with a solution to this problem, there are no constraints in Hmong that could tell us what to do with tones before final voiceless consonants, which never occur in Hmong in final position.

Dupoux and his colleagues (Dupoux et al. 1999; Dehaene-Lambertz et al. 2000), yyy) have proposed that Japanese speakers epenthesize vowels in

borrowed words like ‘srike’ [su.to.rai.ku] for perceptual reasons: they seem to *hear* epenthetic vowels when clusters cannot be parsed into Japanese syllables. Dupoux points out (p.c.) that the Hmong data raise an interesting conondrum for this approach as well. The fact that Hmong deletes sounds in the coda would suggest that speakers don’t perceive them to begin with (else they would epenthesize, as speakers do in Japanese); but the tonal alternations above strongly suggest that Hmong speakers are sensitive to the presence of coda consonants.

Or consider the problems faced by a connectionist model, where loanwords are deformed to match the patterns of native words already stored in the lexicon. No Hmong words have voiceless codas, so there could be no pressure from the lexicon to treat loans with voiceless codas in one way or another.

For these reasons we conclude, albeit with some regret, that the tonal data we collected are modeled just as poorly with constraints (or perception syllables, or connectionism!) as they are with rules.

5. Monosyllabism

As we mentioned above, All Hmong roots and words (except compounds) are monosyllabic. Following a suggestion by McCarthy & Prince, we may model this with a constraint that requires every syllable to be word-initial:

(45) Align-L (syll, word) All syllables are word initial.

The formalism is not as important as the generalization, which would presumably be a part of any adequate analysis of Hmong.

Thus it comes as a rude surprise that polysyllabic roots are freely borrowed both from French and from English. A cursory glance at the preceding data will show how true this is and it flies directly in the face of the alignment constraint immediately above.

We have a technical solution for the problem, but one that does not accord well with native speaker intuitions: viz., that long loanwords are borrowed as compounds. In principle there is a way to test this hypothesis: there is a tone-sandhi rule in Hmong that changes a H tone to a breathy before a falling tone. (PHONG ?) If we had a disyllabic loanword shaped whose syllables were breathy and fall, and if the breathy-toned word were of the shape that normally takes a H tone, we would have evidence for a compound treatment of polysyllabic loans. Conversely, if we found a disyllabic word whose syllables were H and fall we would have evidence

against such an analysis. Unfortunately, we have no H-toned loans in our corpus to check this with, and in fact no breathy tones either.

So we are left with a hypothesis that does not accord well with native speaker intuitions (according to which polysyllabic loans are monomorphemic) and cannot be tested. Similar facts are found in Thai, a monosyllabic language that permits (and encourages) polysyllabic loans in proper names. Thai names like Thanawan Imsuwan, Naruemon Wannapaiboon and Kosum Runglaksameesee are common and, as far as we know, are generally treated as monomorphemic by native speakers.

Polysyllabic loans in Hmong are clearly surprising on any theory of loans that seeks to treat them as nativized forms. Why is this robust grammatical property of Hmong so impotent in the loanword phonology? Again, this is not modeled well with any theory of loanword phonology we know of, rule-based, constraint-based, perception-based, connectionist or otherwise.

6. Directions for future research

We hope to have shown that a constraint-based model is superior in many respects to a model that has only rules, at least insofar as loanwords are concerned. This shows up clearly when we consider how sounds and syllables are borrowed from English and French into Hmong. We used OT, but other constraint-based models should do as well.

But we run into problems modeling other loan-word phenomena that are equally robust. First, Hmong allows 7 distinct tones (H, M, L, rising, falling, breathy, creaky) on its syllables, with no restrictions holding between tone and syllable type. But we find in loans from English that the specific tone a loanword surfaces with is a function of the syllable type of the source word in English. This is hard to model on any theory of loans we are familiar with, including OT. It is our hope that models of tonogenesis may lead us out of this morass, but this is as yet only a hope.

Second, Hmong allows no polysyllabic words other than compounds. But we find no such restriction in the loanword phonology. This can of course be modeled, e.g., using a layered OT approach to that has loanwords subject to a subset of the constraints of a language (Itô & Mester 1995). But we think the modeling does not improve our understanding of the phenomena in this case. Why such a robust property in a language should have no appreciable reflex in loanwords is a mystery to us at present.

We are left with the impression that loanword phonology still poses formidable challenges to models of grammar. Although constraint-based

analyses do substantially better at modeling loanword phonology in some areas (syllables and segments), they fail to provide much insight in others (tones, monosyllabism). We are not optimistic that perception-based or connectionist models will fare any better in these areas, but we leave that issue up to those with the requisite expertise.

References

- Alber, Birgit and Ingo Plag (*to appear*). Epenthesis, deletion and the emergence of the optimal syllable in Croele: the case of Sranan. *Lingua*.
- Andruski, Jean E. and Martha Ratliff (2000). Phonation types in production of phonological tone: the case of Green Mong. *Journal of the IPA* **30** (1/2), 63-82.
- Dupou, Emmanuel, K. Kakehi, Y.. Hirose, C. Pallier and J. Mehler (1999) Epenthetic vowels in Japanese: a perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance*, **25**(6), 1568-1578.
- Dehaene-Lambertz, Ghislaine., E. Dupous and A. Gout (2000). Electrophysiological correlates of phonological processing: a cross-linguistics study. *Journal of Cognitive Neuroscience*, **12**(4), 635-647.
- Golston, C. and P. Yang (2000). The syllable structure of White Hmong. Ms, CSU Fresno.
- Hayes, Bruce (1989). Compensatory lengthening in moraic phonology. *Linguistic Inquiry* **20**, 253-306.
- Heimbach, E. E. (1966). *White Hmong-English Dictionary*. Data paper: number 75. Southeast Asia Program, Cornell University, Ithaca, New York.
- Kehrein, Wolfgang (2000). Phonology without affricates. Ms, Philipps Universität Marburg.
- McCarthy, John J. and Alan S. Prince (1986). Prosodic morphology. Ms., Dept of Linguistics, University of Massachusetts, Amherst, and Program in Linguistics, Brandeis University, Waltham, Mass.
- McCarthy, John J. and Alan S. Prince (1993a). *Prosodic morphology I. constraint interaction and satisfaction*. Ms., University of Massachusetts/Rutgers University.
- McCarthy, John J. and Alan S. Prince (1993b). Generalized alignment. *Yearbook of Morphology* 1993, 79-153.
- Pater, Joe (1999). Austronesian nasal substitution and other NC effects. In Kager, René, Harry van der Hulst and Wim Zonneveld (eds.), *The*

- Prosody Morphology Interface*. Cambridge: Cambridge University Press.
- Prince, Alan. S. and Paul Smolensky (1993). "Optimality theory: constraint interaction in generative grammar". Ms, Rutgers University and University of Colorado, Boulder.
- Silverman, Daniel (1992). Multiple scansion in loanword phonology: evidence from Cantonese. *Phonology* **9**, 289-328.
- Svantesson, Jan-Olaf (1989). Tonogenetic mechanisms in northern Mon-Khmer. *Phonetica* **46**, 60-79.
- Yang, Phong (2000). A reanalysis of tones in White Hmong. Paper presented at WECOL.
- Yip, Moira (1993). Cantonese loanword phonology and Optimality Theory. *Journal of East Asian Linguistics* **2**, 261-291.