

# Tonogenesis

John Kingston

September 14, 2009

Linguistics Department

University of Massachusetts

226 South College, 150 Hicks Way

Amherst, MA 01003-9274

1-413-545-6833, fax -2792

[jkingston@linguist.umass.edu](mailto:jkingston@linguist.umass.edu)

Preliminary draft: Please do not cite or quote without permission.

## 1 Introduction

Tonogenesis<sup>1</sup> raises a host of questions for both phonological and phonetic theory. Among them are:

1. Where do tones come from?
2. What are the features that represent contrasts for phonation and tone?
3. Which of these features are contrastive and which are redundant?
4. Under what conditions can erstwhile redundant features become contrastive?
5. How can one tone develop from a particular source in one language while the opposite tone develops from that same source in another?
6. Is one of these tones phonetically more plausible than the other?

All these questions can be reduced to a single, comprehensive question:

7. How is a phonetic property phonologized?

Tentative answers to these questions and others are developed through a series of case studies, of tonogenesis from consonant phonation contrasts in Yabem and Korean, Kammu and Cham, Vietnamese and Old Chinese, and finally Athabaskan, and of tone splits induced by the phonation contrasts in East and Southeast Asian languages. I also briefly discuss the much rarer case of tone splitting from the phonetic properties of vowels. The chapter ends with discussion of tonogenesis from quite different, non-segmental source in North Germanic and Central Franconian.

---

<sup>1</sup> This chapter owes an enormous debt to David Solnit, with whom I started my study of the historical development of tone more than 20 years ago. Without his insights and compendious knowledge of the history of East and Southeast languages, little of what you see here would have been possible. The few papers and talks we managed to do together – see the references – at best hint at the riches that David brought to our collaboration. This paper is poorer for not still having him as a collaborator and co-author.

## 2 From consonants to tone

### 2.1 Introduction

Tone has developed out of phonation contrasts in adjacent consonants in a large variety of languages and language families. The first two case studies, Yabem and Korean, show how tone can remain redundant on phonation contrasts in preceding consonants, and the next two, Kammu and Cham, shows the transfer of the contrast from preceding phonation contrasts to tone. The next three case studies, of Vietnamese, Old Chinese, and Athabaskan, show how tone evolved from phonation contrasts in following consonants. Finally, the cases of tone splits in East and Southeast Asia show how preceding phonation contrasts can multiply the number of tones on following vowels, i.e. tone splitting rather than tonogenesis. These case studies not only display the variety of phonation contrasts that have induced tonogenesis, but also demonstrate the variety of effects they can have on adjacent vowels' tones.

### 2.2 Yabem and Korean: Contrastive phonation and redundant tone

In Yabem, an Oceanic Austronesian language of the North Huon Gulf (Bradshaw, 1979; Ross, 1993), and Korean (Jun, 1996; Silva, 2006), tone remains redundant on phonation contrasts in preceding consonants.

#### 2.2.1 Yabem

According to Ross (1993), the domain of tone in Yabem is the foot, and two-syllable words consist of a single iambic foot, one-syllable words lack the iamb's initial weak syllable, and three-syllable words consist of an initial strong syllable followed by an iambic foot. Both syllables in a disyllabic foot have the same tone, high or low. Any stops in a high tone foot are also voiceless, but in a low tone foot, they are instead voiced.<sup>2</sup>

---

<sup>2</sup> The only obstruents that are not stops are /s/ and its prenasalized counterpart /<sup>n</sup>s/. Some /s/s are reflexes of Proto-North Huon Gulf (PNHG) \*s and occur in high tone feet, while others descend from PNHG \*z and occur in low tone feet – tone is thus contrastive after /s/ in present-day Yabem. Low but not high tone /s/ can alternate with its pre-nasalized counterpart, /<sup>n</sup>z/. The pre-nasalized fricative is a reflex of \*<sup>n</sup>z and only occurs in irrealis alternants of roots with low tone, see Table 1.

Ross (p. 146) describes tonogenesis in Yabem as the product of tone-voicing harmony, in which the weak syllable of a two-syllable foot acquired the tone and voicing of the strong syllable when the strong syllable began with a stop in Proto-Huon Gulf (PHG), and a strong syllable not beginning with a stop acquired the tone and voicing of the weak syllable when the latter began with a voiced stop in PHG.<sup>3</sup> The alternations in the three realis-irrealis verb paradigms in Table 1 illustrate the harmony of tone and voicing.

	-dèŋ ‘move towards’		-táj ‘weep’		-lu? ‘vomit’	
	realis	irrealis	realis	irrealis	realis	irrealis
1 sg.	gà-dèŋ	yà- <sup>n</sup> dèŋ	ká-táj	yá-táj	gà-lù?	yá-lú?
2 sg.	gò-dèŋ	ò- <sup>n</sup> dèŋ	kó-táj	ó-táj	gò-lù?	ó-lú?
3 sg.	gè-dèŋ	è- <sup>n</sup> dèŋ	ké-táj	é-táj	gè-lù?	é-lú?
1 pl. incl.	dà-dèŋ	dà- <sup>n</sup> dèŋ	tá-táj	tá-táj	tá-lú?	tá-lú?
1 pl. excl., 2 pl.	à-dèŋ	à- <sup>n</sup> dèŋ	á-táj	á-táj	á-lú?	á-lú?
3 pl.	sè-dèŋ	sè- <sup>n</sup> dèŋ	sé-táj	sé-táj	sé-lú?	sé-lú?

Table 1: Realis and irrealis paradigm of Yabem verb stems beginning with a voiced stop *-dèŋ*, a voiceless stop *-táj*, and a sonorant *-lu?*. The irrealis morpheme is prenasalization, which applies to the first voiced stop in the root, if any, and is otherwise unrealized.

The stems *-dèŋ* ‘move toward’ and *-táj* show the covariation of stop voicing and tone within the strong syllable of the foot. No tone is marked on the stem *-lu?* ‘vomit’ because it does not begin with an stop and therefore varies in tone as a function of the voicing of any stop in the prefix, being low when the prefix stop is voiced and low, i.e. after realis singular *gà-*, *gò-*, *gè-*, but high after 1 pl. incl. *tá-* and 3 pl. *sé-* and in all irrealis forms. As this stem is also high after prefixes that do not contain stops, we may infer that high is the default value for tone, and low occurs only in the context of a voiced stop somewhere in the foot. Neither voicing nor its absence can be the default value, as alternations occur in both directions: 1 sg. *gà-* → *ká-*, and 1 pl. incl. *tá-* → *dà-*. These observations, coupled with the existence of a voicing contrast between stops in the Yabem’s pre-tonal ancestor, indicate that voicing is

<sup>3</sup> Complicating this story is the two sources for labial and dorsal voiced obstruents in PHG. First, Proto-Oceanic (PO) \*b, \*g > PHG \*b, \*g > Proto-North Huon Gulf (PNHG) \*b, \*g > Yabem /b, g/. Second, in some morphemes, PO \*p, \*k unpredictably voiced and lenited to PHG \*v, \*ɣ, which in turn disappeared in PNHG yielding zero reflexes in Yabem with low tone.

the source of low tone (as well as of voicing in a weak syllable’s onset), and that high tone and the absence of voicing developed elsewhere.

The feature [voice] is likewise contrastive and tone is redundant synchronically in words containing one or more obstruents. However, there are many words that synchronically contain no stops at all, and yet they contrast for tone (Table 2).<sup>4</sup>

	High Tone		Low Tone	
‘outside’	áwé	àwè <sup>a</sup>	‘woman’	
‘body’	ólí <sup>b</sup>	òlì <sup>c</sup>	‘wages’	
‘mango’	wá <sup>d</sup>	wà <sup>e</sup>	‘crocodile’	
‘prohibition’	yáó	yàò	‘enmity’	
‘hammer (verb)’	-sáʔ	-sàʔ	‘put on top of’	

Table 2: Minimal pairs for tone in words containing no stops. <sup>a</sup> < \*POC (t)apine, <sup>b</sup> < POC \*kultit, <sup>c</sup> < POC \*poli, <sup>d</sup> < POC \*waiwai, <sup>e</sup> < \*puqaya.

As these words contain no voiced stops and the members of each pair are segmentally identical, it’s necessary to recognize that their tone is contrastive rather than predictable. This outcome, where tone remains redundant in a portion of the lexicon but has become contrastive elsewhere as a result of subsequent changes, will recur in other examples discussed in this chapter.

### 2.2.2 Korean

If Yabem exemplifies the dependence of low tone on a [voice] specification in an adjacent stop, then Korean exemplifies the dependence of high tone on [spread glottis] and [constricted glottis] specifications. Jun (1996) shows that following an aspirated [spread glottis] or tense [constricted glottis] stop, the initial L of the accentual phrase’s LHL melody in Seoul and Chonnam Korean is replaced with a high tone, LHL > HHL, while after a laryngeally unspecified lax stop it remains low. Silva (2006) presents similar evidence and also argues that tone is on the verge of becoming contrastive after consonants that are not [constricted

<sup>4</sup> The disappearance of any overt consonantal reflexes of PHG \*v, \*ɣ is presumably the source of low tones in present-day Yabem words that contain no voiced stops.

glottis] as differences in aspiration/voice onset time between [spread glottis] and lax stops are shrinking.

Jun's instrumental evidence shows that F0 remains higher through the entire first vowel of accentual phrases following [spread glottis] or [constricted glottis] consonants than following lax stops and that these F0 differences are far larger and last far longer than those following stops contrasting for [voice] in her English and French data. Elsewhere in the accentual phrase in Korean, however, F0 is elevated less, over a shorter portion of the vowel following these consonants. Rather than being differences in tone, these smaller, shorter F0 differences in non-initial position are described as "phonetic perturbations" of the vowel's F0 or as "microprosodies" in contrast to the "macroprosodies" consisting of the tones that make up the accentual phrase melody. Despite the perturbations' or microprosodies' smaller size and extent, they are undoubtedly the historical source of the much larger tone difference at the beginnings of accentual phrases, where they have been exaggerated, presumably to better demarcate the beginning of that prosodic constituent. Despite the exaggerated size and extent of the difference between tones at the beginnings of accentual phrases, for now they remain just as predictable from/redundant on the phonation specification of the preceding consonant as do the low versus high tones in Yabem words containing stops (but cf. Silva, 2006).

The Yabem and Korean examples represent the first step toward the phonologization of consonantal perturbations of F0 in adjacent vowels as tone. The F0 differences remain predictable from the phonation differences in the consonants, yet they are exaggerated in both size and extent in both languages, so they cannot still be described as mere phonetic perturbations. Even so, they are also not yet contrastive, except in those Yabem examples where the consonant that was source of tone is no longer present or its phonation is no longer distinctive.

### 2.3 Kammu and Cham: Contrastive tone

Kammu and Cham illustrate the next step in the historical development of tone contrasts from phonation contrasts in preceding consonants. In both cases, the languages were previously toneless, and their close relatives remain so.

#### 2.3.1 Kammu

Table 3 shows that in two of the western dialects of the Mon-Khmer language, Kammu, a low versus high tone contrast has replaced a voiced versus voiceless contrast in the preceding consonants, which is still preserved in the eastern dialect of this language on both stops and sonorants (Svantesson, 1989; Suwilai, 2003).

On long vowels, a rising-falling tone in the first tonal western dialect corresponds to a voiceless initial in the non-tonal eastern dialect, while a simple high tone corresponds to that initial in the second tonal western dialect. In both tonal dialects, a low tone corresponds to a voiced initial in the non-tonal dialect. In the second tonal dialect, voiceless aspirated stops correspond to voiced stops in the eastern dialect, while in the the first tonal dialect, voiceless unaspirated stops do. A phonation contrast is thus maintained in a shifted form in stop-initial words in the second tonal dialect, while it is completely neutralized or more precisely *transferred* to tone in the first tonal dialect's cognates of these words. Contrast transfer is also referred to as “transphonologization” (Hyman, 2008)/ In the register dialect, a clear or tense voice quality and high tone corresponds to voiceless initials in the non-tonal dialect, while a breathy voice quality and low pitch corresponds to voiced initials.<sup>5</sup> In all three western dialects, the phonation contrast has been completely lost in sonorants, and it

---

<sup>5</sup> “Register” refers to a cluster of phonetic differences, breathy or lax versus tense, clear, or modal voice, low versus high F0 or pitch, and higher versus lower vowel qualities that may occur singly or in combination in so-called “register” languages (Huffman, 1976; Gregerson, 1976; Thongkum, 1987; Denning, 1989). These properties may all be concomitants of expanding the pharynx by advancing the tongue root and lowering the larynx (Lindau, 1979; Edmondson & Esling, 2006). In many Mon-Khmer languages, the first property in each pair appears in words that began with voiced consonants (obstruents and sonorants) in an earlier form of the language, while the second property appears in words that began with voiceless consonants. Outside of the Vietic and the Kammu languages discussed here, any F0 differences that accompanied the other differences between the two registers have not become contrastive; instead voice quality and in some instances the differences in vowel quality have.

Gloss	E. Kammu	W. Kammu Tone 1	W. Kammu Tone 2	W. Kammu Register
‘rice wine’	bu:c	pù:c	p <sup>h</sup> ù:c	pu:c
‘to take off clothes’	pu:c	pû:c	pú:c	pû:c
‘to cut down a tree’	bok	pòk	p <sup>h</sup> òk	pok
‘to take a bite’	pok	pók	pók	pók
‘to chew’	bu:m	pù:m	p <sup>h</sup> ù:m	pu:m
‘to fart’	pu:m	pû:m	pú:m	pû:m
‘stone’	gla:ŋ	klà:ŋ	k <sup>h</sup> là:ŋ	kla:ŋ
‘eagle’	kla:ŋ	klâ:ŋ	klá:ŋ	klâ:ŋ
‘to weigh’	ɟaŋ	càŋ	c <sup>h</sup> àŋ	caŋ
‘astringent’	caŋ	câŋ	cáŋ	câŋ
‘to fear’	ŋɔʔ	ŋòʔ	ŋòʔ	ŋɔʔ
‘paddy rice’	<sup>h</sup> ŋɔʔ	ŋóʔ	ŋóʔ	ŋóʔ
‘flower’	ra:ŋ	rà:ŋ	rà:ŋ	ra:ŋ
‘tooth’	<sup>h</sup> ra:ŋ	râ:ŋ	rá:ŋ	râ:ŋ
‘to chase’	waʔ	wàʔ	wàʔ	waʔ
‘monkey’	<sup>h</sup> waʔ	wáʔ	wáʔ	wáʔ

Table 3: Correspondences of tone and register in three western dialects of Kammu to the [voice] contrast on initial stops (above the double line) and sonorants (below the double line) in the eastern dialect. Data are from Suwilai (2003). The toneless E. Kammu apparently corresponds to Svantesson’s 1989 “Southern” Kammu, and the first tonal W. Kammu dialect apparently corresponds to his “Northern Kammu.” The register dialect of W. Kammu appears to be very similar to Svantesson’s Lamet, which also developed register rather than tone from the original [voice] contrast. The circumflex ‘â’ indicates a rising-falling tone, the acute accent ‘á’ a high tone, and the grave accent ‘à’ a low tone. The ‘<sup>h</sup>’ preceding a sonorant symbol indicates that it’s voiceless, the ‘<sub>a</sub>’ beneath a vowel indicates breathy voicing.

has been replaced by a contrast for tone or register.

Following other classes of consonants, tone and register are entirely predictable: tone being high or rising-falling after voiceless aspirated stops, /s/, /h/, i.e. [spread glottis] consonants, and after glottalized glides /<sup>ʔ</sup>j, <sup>ʔ</sup>w/, and implosives /ɓ, ɗ/ – the implosives correspond to glottalized nasals /<sup>ʔ</sup>m, <sup>ʔ</sup>n/ in the non-tonal dialect, so this last source of high or rising-falling tones is the class of glottalized or [constricted glottis] sonorants. Surprisingly, the tone is low following /ʔ/ itself. In Lamet, a closely related language that developed a

register contrast much like that in the W. Kammu register dialect in Table 3, creaky voice appears after all the same classes of consonants as high tone in the two tonal dialects, but also after /ʔ/, while breathy voice appears otherwise where low tone does in the tonal dialects.

The development of tone or register in the W. Kammu dialects also shows that it may be the phonetic realization rather than phonological specification of a consonant that determines which tone or register develops after it. Voiced sonorants are generally not thought to be specified for [voice], yet they pattern with the voiced stops, which are specified for this feature. Could the voiced sonorants actually be specified for [voice] in the ancestor of these languages and to this day in E. Kammu? After all, they contrast with voiceless sonorants in that ancestor and E. Kammu. This contrast is, however, more likely to be for [spread glottis], with the voiceless sonorants being specified and the voiced ones unspecified for both [spread glottis] and [voice]. Similarly, voiceless fricatives such as /s/ are not generally treated as specified for [spread glottis], even though they are pronounced with the glottis opened wide. They also pattern in tonogenesis with voiceless aspirated stops and /h/, which are specified [spread glottis]. As all these segments also pattern together with the voiceless sonorants, here and in other examples presented below, they could all be specified for [spread glottis]. Alternatively, voiced sonorants are certainly produced with voicing, and likewise voiceless fricatives are produced with a wide-open glottis (Löfqvist & Yoshioka, 1980; Yoshioka, Löfqvist, & Hirose, 1981), so they phonetically resemble the segments specified for these features, and that phonetic resemblance could be enough for them to pattern with the segments specified for [voice] or [spread glottis] in these sound changes.

### 2.3.2 Cham

Two Chamic languages, East Cham (Phu, Edmondson, & Gregerson, 1992) and Utsat (Maddieson & Pang, 1993) have also developed tone from preceding consonants (see also Thurgood, 1999).

E. Cham contrasts low with non-low<sup>6</sup> tone in syllables that began in Proto-Cham with voiced versus voiceless stops<sup>7</sup> – W. Cham has a register contrast instead, where higher vowels with a laxer or breathier voice quality contrast with lower vowels with a tenser voice quality (Table 4a versus b). In both E. and W. Cham, the voice contrast has been completely lost in initials. The low tone is low level in open syllables, rising in syllables with a long vowel ending in a [ʔ], and rising-falling in syllables with a short vowel ending in a [ʔ], while the non-low tone is high level in open syllables and those with a long vowel ending in a [ʔ] and high and slightly falling in those with a short vowel ending in a [ʔ] – the final [ʔ]s are the reflexes of all earlier final stops. The low tone in E. Cham syllables not ending in stops is also accompanied by a breathier voice quality than co-occurs with the non-low tone in those syllables. Voice quality does not differ systematically between low and non-low tone syllables ending in [ʔ].

For Proto-Cham syllables beginning with a sonorant and not ending in a stop (Table 4c), W. Cham unexpectedly shows the breathy register, but in those ending in a stop it shows the expected tense register instead (Table 4d). Both kinds of syllables bear the non-low tone in E. Cham. The unexpected appearance of breathy register in these W. Cham forms may have arisen in the same way as the low tone following voiced sonorants in the tonal varieties of W. Kammu: because the sonorants are phonetically voiced they behave like voiced stops, even though they are not specified for [voice]. Unlike in the ancestor of the W. Kammu dialects, sonorants did not contrast for [voice] in Proto-Cham, so there was no phonological encouragement for the voiced sonorants to pattern with the voiced stops, and they must instead have done so because they were phonetically similar. This extension of the breathy register would have been inhibited by a final [ʔ] because its glottal constriction would be incompatible with the glottal spreading that would produce breathy voice.

Utsat distinguishes five tones, high level (55), mid level (33), low level (11), rising (24),

---

<sup>6</sup> “Non-low” is the descriptor used by Phu et al. (1992).

<sup>7</sup> The voicing contrast is maintained in other Chamic languages not illustrated here; they are the sources for the reconstructions in Table 4 below.

Gloss	Proto-Cham	W. Cham	E. Cham
(a) Proto-Cham initial voiced stops			
‘give’	*brɛj	prɔːj	prɛːj
‘fill, full’	*bap	pauːʔ	pàːʔ
‘two’	*dua	tɔa	twàː
‘to sit, live, stay’	*dɔːk	tɔːʔ	tòːʔ
‘forest, wild savage’	*glay	klaj	-klàːj
‘kettle, glazed clay pot’	*gɔʔ	kɔʔ	kòʔ
(b) Proto-Cham initial voiceless stops			
‘master’	*pɔ	pɔː	póː
‘four’	*paːt	paːʔ	páːʔ
‘to follow’	*tuj	tui	túːj
‘mountain range’	*cət	cəːʔ	cáːʔ
‘monkey’	*kra	kraː	kráː
‘white’	*kɔːʔ	kɔːʔ	kóːʔ
(c) Proto-Cham initial sonorants in syllables not ending in stops			
‘to forget’	*wər	wɔr	wár
‘to go, walk’	*naw	naːw	náːw
‘insect’	*ruaj	ruaj	róːj
(d) Proto-Cham initial sonorants in syllables ending in stops			
‘to take, fetch, get’	*mat	məːʔ	míʔ
‘grass, weeds’	*rək	rəːʔ	(haː)róʔ
‘to fall into’	*lɛʔ	lɛʔ	léʔ

Table 4: Reconstructed Proto-Cham forms and their W. and E. Cham reflexes for words that begin with (a) voiced stops, (b) voiceless stops, (c) sonorants in syllables not ending in stops, and (d) sonorants in syllables ending in stops.

and falling (42).<sup>8</sup> The mid tone occurs in both open and ʔ-final syllables, the low and high tones only in open syllables, and the rising and falling tones only in ʔ-final syllables – there are exceptions to all these generalizations. The high tone appears in syllables that ended in /h/ and in some instances /s/ in Proto-Cham (Table 5a), regardless of whether the Proto-Chamic source of the initial was a voiced stop. Low tone appears in what were open syllables beginning with a voiced stop in Proto-Cham (Table 5b), and mid tone in syllables

<sup>8</sup> The numbers refer to Chao’s five-level scale for notating tones’ relative pitch values, where 1 is the lowest level and 5 the highest. A level tone is indicated by two identical numbers, a contour by different numbers that identify its initial and final levels. These numbers will be displayed after the syllable bearing the tone to represent tones in languages where the conventional diacritics do not provide enough information about the tones’ pronunciations.

with neither of these properties in Proto-Cham (Table 5c), that is, open syllables beginning with a voiceless stop or fricative or a nasal.<sup>9</sup> If the Proto-Cham syllable instead ended in a stop, the rising tone developed with these initials (Table 5d), while the falling tone developed in stop-final syllables when the initial was a voiced stop (Table 5e). The Proto-Cham voiced stops became voiceless aspirated in Utsat.

With one exception, the tonal developments in Utsat are quite similar to those in E. Cham: low tones develop in syllables that began with voiced stops in Proto-Cham and a higher tone in other syllables, and the tones' F0 trajectories vary allophonically as a function of whether the syllable ends in a [ʔ] in the present-day language. If it does, then a contour tone develops, but if it does not, then a level tone does. Because the Proto-Cham voiced stops have become voiceless aspirated stops in Utsat, while the Proto-Cham voiceless stops remained unaspirated, the difference in F0 level remains predictable from how the syllable begins, just as the difference in F0 trajectory remains predictable from how it ends.

The exception is the high tone in Utsat, which developed in syllables ending in /h/ or /s/ in Proto-Cham. This development has no correspondent in E. Cham. This tone is also exceptional in that it develops regardless of whether the initial was a voiced stop in Proto-Cham; that is, its development is entirely independent of the system of register contrasts. The practical effect of this development is that high tones contrast with mid or low tones in Utsat because high tones appear in words with all initials.

## 2.4 Interim discussion

In Yabem, W. Kammu, E. Cham, and Utsat, low tones develop on vowels following voiced stops and non-low tones elsewhere. These developments suggest that something about pronouncing a voiced stop also automatically lowers F0 in following vowels. That is, these tones

---

<sup>9</sup> The word *nãʔ ʒʒ* 'six' has a mid tone on a stop-final syllable, but there is no final stop in its Proto-Cham source, *\*nam*. Moreover, not a single word reconstructed with a final stop in Thurgood's (1999) list of Proto-Cham forms has a Utsat reflex with a mid tone. They all instead have either a rising or falling tone. Final stops in any Utsat words with mid tones must therefore be secondary developments. The generalization would then be that mid tone appears on open syllables not beginning with a voiced stop in Proto-Cham.

Gloss	Proto-Cham	Utsat
(a) Syllables ending in /h/ > High tone 55		
‘thin (material)’	*(li)pih	pi 55
‘all, finished, done’	*(ʔa)bih	p <sup>h</sup> i 55 p <sup>h</sup> i 55
‘ten’	*pluh	piu 55
‘fruit, egg, classifier’	*bɔh	p <sup>h</sup> o 55
‘land, earth, soil’	*(ta)nah	na 55
(b) Open syllables beginning with voiced stops > Low tone 11		
‘buffalo’	(ku)bau	phau 11
‘two’	*dua	thua 11
‘tooth’	*gigɛj	k <sup>h</sup> ai 11
(c) Other open syllables > Mid tone 33		
‘dream’	*(lu)pəi	pai 33
‘eye’	*(ma)ta	ta 33
‘male’	*(la)kəi	kai 33
‘one’	*sa	sa 33
‘hand, arm’	*(ta)ŋa:n	ŋa:n 33
‘bone’	*(tu)la:ŋ	la:ŋ 33
‘fly, bug, insect’	*ruaj	zuai 33
(d) Stop-final syllables not beginning with a voiced stop > Rising tone 24		
‘four’	*pa:t	paʔ 24
‘boil, cook’	*(sa)tuk	tuʔ 24
‘mountain range’	*cət	tsəʔ 24
‘sick, painful’	*(sa)kit	kiʔ 24
‘ripe, cooked’	*(ta)saʔ	saʔ 24
‘child’	*(ʔa)na:k	naʔ 24
(e) Stop-final syllables beginning with a voiced stop > Falling tone 42		
‘to do, work’	*buat	phuaʔ 42
‘to sit, live, stay’	*dɔ:k	t <sup>h</sup> oʔ 42
‘to step on, tread’	*ʃɛʔ	seʔ 42

Table 5: Proto-Cham words and their Utsat reflexes, where the Proto-Cham words consist of (a) syllables ending in /h/, (b) open syllables beginning with voiced stops, (c) open syllables not beginning with voiced stops, (d) syllables ending in stops and not beginning with voiced stops, or (e) syllables ending in stops and beginning with voiced stops. The parenthesized portions of the Proto-Cham words have no reflexes in the present-day Utsat forms.

arise through exaggeration of a phonetic side effect of producing voicing in a stop or in obstruents generally. Most attempts to find the articulation that has this side effect have been directed at articulations the speaker produces to resolve the aerodynamic conflict be-

tween keeping voicing going during the stop closure and producing the characteristic noise burst at its release. On the one hand, the stop closure is intended to trap air inside the oral cavity, such that pressure builds up inside it. The rise in intraoral air pressure ( $P_O$ ) then produces a brief but intense burst of noise when the stop is released as the trapped air rushes forcefully out of the mouth. On the other hand, voicing depends on a continued flow of air up through the glottis, which itself depends on  $P_O$  remaining lower than subglottal air pressure ( $P_S$ ). The rapid rise in  $P_O$  behind the stop closure downstream would soon reduce the pressure drop across the glottis below the threshold for maintaining voicing if speaker did not act to slow its rise. The search has been for an articulation that both maintains the pressure drop across the glottis on which voicing depends and that also automatically lowers the fundamental frequency (F0) of adjacent vowels (Hombert, 1978; Hombert, Ohala, & Ewan, 1979).<sup>10</sup>

Perhaps, the most plausible candidate for such an articulation is larynx lowering, which slows the rise in  $P_O$  by enlarging the oral cavity<sup>11</sup> and at the same time slackens the vocal folds by tilting the cricoid cartilage forward relative to the thyroid cartilage (Honda, Hirai, Masaki, & Shimada, 1999).<sup>12</sup> The F0 lowering that results from this slackening is considered to be in an automatic, indeed accidental mechanical consequence of larynx lowering because the purpose of this articulation in voiced stop is to expand the oral cavity and keep the rise

---

<sup>10</sup> Löfqvist, Baer, McGarr, & Seider Story (1989) report higher levels of cricothyroid activity preceding voiceless than voiced stop closures. As this muscle's contraction stretches the vocal folds by tilting and sliding the thyroid cartilage forward, these authors propose that its higher level of activity is intended to aid glottal abduction in turning voicing off. If this muscle remains contracted long enough it might also raise F0 in vowels following voiceless stops. However, what is needed for the languages discussed here in which low tone has developed after voiced stops is not an active devoicing articulation that also raises F0 in adjacent vowels, but instead an active voicing articulation that also lowers F0.

<sup>11</sup> Larynx lowering is only one of a number of manoeuvres that actively expand the oral cavity; others include tongue root advancement, tongue body lowering, and soft palate raising (Bell-Berti & Hirose, 1975; Westbury, 1983). The cheeks, lips, soft palate, and other compressible tissue lining the oral cavity are also distended by the rising  $P_O$ , which expands the oral cavity passively (Ohala & Riordan, 1979).

<sup>12</sup> Hombert et al. (1979) had themselves suggested that larynx lowering was the most plausible candidate, but until Honda et al. demonstrated that the curvature of the cervical spine caused the cricoid cartilage to tilt forward relative to the thyroid cartilage as the larynx is pulled down, there was no convincing exposition of the mechanism whereby larynx lowering would lower F0. Forward tilt of the cricoid cartilage would shorten and slacken the vocal folds and thereby cause them to vibrate slower.

in  $P_o$  in check. However, Honda et al. (1999) as well as Collier (1975), Ewan (1976), and Erickson, Honda, Hirai, Beckman, & Seiji (1994) all show that larynx lowering is also one of the manoeuvres speakers use when they wish to lower F0 deliberately. Their results raise the possibility that larynx lowering in a voiced obstruent is actually intended to lower F0 as well as  $P_o$ .

Why would a speaker wish to lower F0 when producing a voiced obstruent? Kingston & Diehl (1994, 1995), and Kingston, Diehl, Kirk, & Castleman (2008) argue that a low F0 in vowels flanking a voiced stop integrates perceptually with the presence of low frequency periodic energy, namely, voicing, in the stop closure itself to enhance the percept of the presence of low frequency energy in and near the stop (see also Stevens & Blumstein, 1981), and that it is the percept of low frequency energy rather than of either voicing or low F0 that conveys to the listener that the stop is [voice]. Integration unifies the phonation of the stop and the pitch of an adjacent vowel phonetically, at stage in perception in between the raw acoustics of the signal and the listener's recognition of the stop's value for an abstract distinctive feature [voice]. This line of reasoning suggests that the lower F0 next to voiced obstruents may not be an automatic side effect of an manoeuvre whose purpose is to manipulate  $P_o$ , but instead a product of a deliberate manoeuvre that is intended by the speaker to lower F0. Larynx lowering is particularly interesting in this regard, because it may serve both purposes at once.

This account extends naturally to the register differences that have developed from the [voice] contrast in stops in W. Kammu, Lamet, W. Cham, and elsewhere (see also Henderson, 1967; Matisoff, 1973; Denning, 1989; Thurgood, 2002, for recognition of this linkage). In breathy voice, there is far more energy in the first than higher harmonics than in modal or tense voice, and in higher vowels, the first formant (F1) is lower than in lower vowels. While both differences can be interpreted as automatic concomitants of producing voicing in an adjacent stop, their exaggeration and phonologization in the development of register contrasts in these languages can equally easily be construed as an attempt to strengthen the

percept of low frequency energy next to the stop. Like the development of tone contrasts, the low frequency property becomes distinctive as a register contrast once the stop itself no longer differs in voicing. The essential difference between the two lines of development is that tonogenesis selects the differences in F0 while registrogenesis selects the differences in voice and vowel quality.

Implicit in this account is the claim that voicing during the consonantal constriction, low F0 in the adjacent vowels, breathy or lax voice, and lower F1 are all acoustic evidence that the stop is specified [voice], and the absence of these properties is evidence that it's not specified for this feature (or alternatively as evidence for [+voice] versus [-voice] specification if this feature is not privative). In principle, none of these properties is perceptually more important than any other, and therefore none of the articulations that produce them is any more likely to be controlled or automatic than the articulations that produce the others. [voice] is therefore an abstract feature that can be realized phonetically in a variety of ways. The contrast would remain for [voice] until the language reaches a stage when the consonants no longer differ phonetically from one another in any context, and all the phonetic differences are those produced during the vowel. Only then would the phonetic perturbations of the vowel's F0, voice quality, or vowel quality have been phonologized as and transferred to tone or register contrasts.

Halle & Stevens (1971) propose a different unification, which replaces the representation of both the [voice] contrast in obstruents and tone contrasts in vowels with the features [stiff] and [slack],<sup>13</sup> which represent differences in the tension applied to the vocal folds (see Avery & Idsardi, 2001, for a dramatic update and expansion of Halle & Stevens's proposals) and thereby giving primacy to differences in vocal fold tension that were thought to be automatic side effects of other articulations, with other purposes. A stop specified for [slack] would be

---

<sup>13</sup> Treating these contrasts as privative is anachronistic, as Halle & Stevens proposed binary features, [±stiff] and [±slack]. Nonetheless, nothing in their proposal depends on this difference, given their acknowledgment that the combination [+stiff, +slack] does not and cannot occur. Moreover, the occurring combination [-stiff, -slack] is formally equivalent to the absence of specification for both features if they are instead privative.

pronounced with voicing, while one specified for [stiff] would be pronounced without it. Halle & Stevens propose that these adjustments of vocal fold tension would contribute to bringing about voicing or its absence, and that voicing in obstruents is thus a side effect of the vocal folds being slack enough. Vowels specified for [slack] and [stiff] would be pronounced with a low and high tones, respectively.

If, as their names imply, [stiff] and [slack] refer narrowly to the tension applied to the vocal folds, then these features do not encompass the conditions required to maintain voicing in a stop. The challenge is to keep  $P_O$  enough below  $P_S$  that air continues to flow up through the glottis. Slackening the folds both helps and hinders meeting that challenge: slacker folds will vibrate for a smaller pressure drop across the glottis than stiffer ones, but slacker folds also let a larger volume of air pass through the glottis per unit of time and thereby accelerate the rise in  $P_O$ . Speakers apparently rely instead on slowing the rise in  $P_O$  by expanding the oral cavity, actively and passively. In this narrow interpretation, Halle & Stevens's features do not abstract away from the phonetic realization of the phonation and tone contrasts, so much as they subsume them under a single contrast, whose phonetic realization differs systematically as a function of the segment's value for other features. In this respect, too, their proposal differs from Kingston & Diehl's (1994) proposal.

I will suggest below that the features [stiff] and [slack] remain useful but only so long as they are independent of [voice], and for that matter the other two features needed to represent laryngeal contrasts in consonants, [spread glottis] and [constricted glottis].

The Korean facts show that obstruents with other phonation types than [voice] raise rather than lower  $F_0$ . Jun (1996) attributes this difference to the [spread glottis] and [constricted glottis] stops also being [stiff], which is apparently a byproduct of higher levels of thyroarytenoid (AKA vocalis) contraction in these stops compared to the lax stops – cricothyroid activity does not differ between the three classes of stops (Hirose, Lee, & Ushijima, 1974; Hirose, Park, Yoshioka, Sawashima, & Umeda, 1981). The differences in the amount and timing of thyroarytenoid contraction between these stops appear to be intended to con-

tribute to differences between them in the size and timing of the glottal opening (see also Kagaya, 1974) and in the [constricted glottis] stops to the firmness with which the vocal folds are pressed against one another, but they clearly affect F0, too.

In all these examples, tonal contrasts have emerged on vowels under the influence of the laryngeal articulations of preceding consonants. I turn my attention next to tonogenesis from the laryngeal articulations of following consonants. I begin with an account of the probable origin of tones from this source in the ancestor of Vietnamese and in Old Chinese and continue with an account of tonogenesis from following consonants in Athabaskan.

## 2.5 Vietnamese and Old Chinese: Tonogenesis from following consonants

### 2.5.1 Vietnamese, with an excursus on Tamang

According to Haudricourt (1954b), the tone contrasts of present-day Vietnamese developed in two stages. First, a final stop induced a rising tone, a final voiceless fricative a falling tone, and a level tone developed in other syllables, which were either open or ended in a nasal (columns in Table 6). A considerable time later, each of these tones then split in two under the influence of a [voice] contrast in syllable initial stops, with words beginning with voiced stops developing lower pitches than those beginning with voiceless ones (rows in Table 6). The result is the six-way tone contrast displayed in Table 6, which can be observed in Vietnamese spoken in the northern part of Vietnam.

		Following Consonants		
		CV, CVN	CVK, CVʔ	CVs, CVh
Preceding consonants	Voiceless	*pa > pa high level <i>ngang</i>	*pak > pak high rising <i>sac</i>	pas > pa high falling <i>hoi</i>
	Voiced	*ba > pa low level <i>huyen</i>	bak > pak low rising <i>nang</i>	bas > pa low falling <i>nga</i>

Table 6: Tonogenesis in Vietnamese, first from following consonants, in any, later from preceding consonants (“k” = a following stop). The italicized words are Vietnamese names for each of the tones.

Thurgood (2002) argues that the tones attributed to final and initial stops in Vietnamese did not arise directly from the stops themselves but instead from differences in voice quality that perturbed F<sub>0</sub>. The initial distinction between a rising and level tone arose from a contrast between creaky and modal voice in syllables not ending in stops in Proto-Vietic (Diffloth, 1989). Tones in syllables ending in stops subsequently merged with those arising in creaky voice syllables because final stops were pronounced with a simultaneous glottal closure that would also induce creaky voice on preceding vowels. Similarly, the subsequent tone split arose from an earlier difference between a breathy and a modal voice quality induced by preceding voiced and voiceless stops and not directly from the [voice] contrast itself. Only the falling tones that developed in syllables that ended in voiceless fricatives remain attributed to the consonants themselves (p. 336).

Although Thurgood emphasizes that differences in voice quality, creaky versus modal and breathy versus modal, are the immediate sources of the rising versus level and low versus high tonal contrasts in Vietnamese, he also recognizes that the ultimate causes are differences in the properties of consonants, except, of course, for the original creaky versus modal contrast in open syllables. Thurgood also asserts that it is specifically the consonants' laryngeal properties that matter, the constricted glottis of the simultaneous glottal closure in final stops, the spread glottis of the final voiceless fricatives, and the voicing of initial stops.

Before turning to Chinese, it is worthwhile to consider the historical development of tone and voice quality in the Tamang languages and dialects, as they indicate that F<sub>0</sub> and voice quality differences can covary, but without one necessarily emerging before and serving as the source for the other. The Tamang languages are a group of closely related Tibeto-Burman languages or dialects spoken in Nepal, in which tone, voice quality to a lesser extent voicing in preceding obstruents covary (Mazaudon, 1978, 2009; Mazaudon & Michaud, 2008). Mazaudon reconstructs two tones in the ancestor of Tamang, each of which was split in two under the influence of the laryngeal articulations of preceding consonants. The lower reflexes

of the split appear in syllables that originally began with voiced obstruents (voiced stops and affricates and \*z) and sonorants (nasals, liquids, and glides), the higher reflexes in syllables that began with voiceless obstruents (voiceless unaspirated and aspirated stops affricates and \*s) and sonorants. The voiced stops subsequently merged with the voiceless unaspirated stops, and the voiceless sonorants merged with their voiced counterparts.

In five of eight languages (Risiangku and Sahu Tamang, Tukche and Syang Thakali, and Ghachok Gurung), the lower two tones that developed after earlier voiced obstruents and sonorants are pronounced with breathy voicing, while the higher series that developed after earlier voiceless obstruents and sonorants are pronounced with modal voicing.

Mazaudon & Michaud's (2008) instrumental study of Risiangku Tamang shows that word-initial preceding stops are variably voiced in syllables bearing the two lower tones, i.e. in the syllables where the etymological source of the lower tones was voicing in the stops. However, voicing more often continued through part than all of the stop closure, none of the five speakers in this study produced these stops consistently with voicing, and the five speakers differed in how frequently they did so, from as few as 10% of the tokens to perhaps as many as 60%. The highest frequency is observed for just one speaker in one tone; otherwise, the frequency of voiced pronunciations is less than 40% and typically no more than 20%. Although one could treat these occasional and partially voiced pronunciations as traces of the etymological sources of the two lower tones, a better synchronic account would instead treat them as inconsistent and infrequent covariants of the consistently low F0 and breathy voice of these syllables.

The remaining three Tamang languages have undergone further changes in tone and voice quality. In Taglung Tamang and Marpha Thakali, one of the lower tones has become high falling (51) and lost its breathy voice quality, and both lower tones have been raised and lost their breathy voice quality in Ngawla Manang. This last language exhibits one last, striking development: in syllables that had one of the two lower tones, breathiness has been transferred to preceding stops in the form of aspiration. The sequence of sound changes

affecting the preceding consonant in this tone was thus  $**ba > *p\grave{a} > p^ha$ .

These facts show that differences in F0 and voice quality act in concert to distinguish morphemes from one another in five conservative Tamang languages, and that the voice quality difference has been subsequently been lost in the three more innovative ones.

### 2.5.2 Old Chinese

Middle Chinese, as documented in the period between the fourth and sixth century CE, distinguished three tones on syllables not ending in stops, *ping* ‘level’ (A), *shang* ‘rising’ (B), *qu* ‘departing’ (C); a fourth tone occurred on syllables ending in stops, *ru* ‘entering’ (D) (Sagart, 1999). According to Sagart, tones A, B, and D can be extrapolated back a millenium to Old Chinese,<sup>14</sup> while many instances of Middle Chinese tone C were derived morphologically from tones A or B (Downer, 1959; Mei, 1980; Schuessler, 1985). Tone A arose in Old Chinese in open syllables and syllables ending in sonorants (glides and nasals), and tone B in syllables ending in a glottal stop (Mei, 1970). The instances of tone C that did not arise through derivation arose from a final /h/, which itself was the reflex of an earlier final /s/ (Haudricourt, 1954a). Both the etymological glottal stop that was the source of tone B and the /h/ that was the source of underived tone C were lost by the Middle Chinese period if not well before. Since the Middle Chinese period, tones have split in the modern Chinese languages under the influence of the [voice] contrast in preceding stops, with lower reflexes appearing in syllables that began with voiced consonants than in those that began with voiceless ones – voiced sonorants sometimes pattern with the voiced stops, other times differently, and the voiceless aspirated and unaspirated stops generally pattern together.

Tonogenesis in Old Chinese and the subsequent splitting of these tones in developments from Middle Chinese to the modern languages are thus strikingly similar to the history of tones in Vietnamese. Tones first emerge under the influence of the laryngeal articulations

---

<sup>14</sup> Haudricourt (1954a) and Pulleyblank (1962) propose that Old Chinese did not distinguish syllables for tone, and that the emergence of the four tones characterizes the development of Middle from Old Chinese in the 6th-7th centuries CE. Haudricourt (1954b) proposes that Old Vietnamese was also toneless until the 6th century CE, when it, too, developed four tones, under the influence of Chinese.

of following consonants and later multiply through splitting under the influence of preceding consonants. The laryngeal articulations of following consonants that first produce tones are constricting and spreading the glottis, while the laryngeal articulation of preceding consonants that later splits the tones is voicing. The parallelism is reinforced by the tonal adaptation of the earliest Chinese loans into Vietnamese at the beginning of the Common Era. The loans were borrowed into the Vietnamese tone classes that developed from the same following laryngeal articulations: tone B loans into the rising tone class – both arose from syllable-final glottal constriction (middle column of Table 6) – and tone C into the falling tone class – both arose from syllable-final glottal spreading (right hand column in Table 6). Finally, there is even parallelism in what appear to be exceptional instances of Chinese tone B and of the Vietnamese rising tone class in etyma which lack the requisite syllable-final [constricted glottis] consonant yet still show up with these tones. Rather than reconstructing an earlier voice quality contrast as Diffloth (1989) and Thurgood (2002) do for Vietnamese, Sagart (1999) instead reconstructs a series of final voiced stops that developed into glottalized nasals before inducing tone B in Old Chinese. Glottalization was subsequently lost, leaving tone B behind on words ending in nasals. Despite these striking parallels, the two instances of tonogenesis must be independent because tones emerged in Old Chinese from the laryngeal articulations of following consonants centuries before they did in Proto-Vietic.<sup>15</sup>

Now that tonogenesis from constricted and spread glottis articulations of syllable-final consonants in Vietnamese and Old Chinese has been sketched, we can take up tonogenesis in Athabaskan in some more detail.

---

<sup>15</sup> This last observation obviously depends on Sagart (1999) being correct in pushing the origin of the four Middle Chinese tones back to Old Chinese, as opposed to Haudricourt (1954a) and Pulleyblank (1962), who treat Old Chinese as not yet distinguishing syllables for tone and Haudricourt (1954b) who treats Vietnamese as having developed tone contemporaneously with its development in Middle Chinese in the 6th century CE.

## 2.6 Athabaskan

Many of the Athabaskan languages of northwestern Canada (e.g. Chipewyan and Gwich'in), as well as the Apachean subgroup (including Navajo) in the southwestern United States contrast morphemes for tone, while the Athabaskan languages of southwestern Alaska (e.g. Ahtna) and the Pacific Coast of Oregon and northern California (e.g. Hupa) do not (Leer, 1979, 1999, 2001; Krauss, 2005; Kingston, 2005). Most of the non-tonal languages retain a contrast between glottalic and non-glottalic stops at the ends of stems that has been replaced by tones in the tonal languages.

### 2.6.1 Stem rime contents

Fully five contrasts between the segments that composed stem rimes in the proto-language interacted in determining how tone evolved, three between stem-final consonants (1-3) and two between nuclei (4-5):<sup>16</sup>

1. Glottalic vs non-glottalic consonants (C' or 'C versus C),
2. Stops vs sonorants (K versus R),
3. Stops vs fricatives (K versus X),
4. Full vs reduced (or long vs short) nuclei (VV versus V), and
5. Full vowel nuclei ending with glottal constriction, a 'glottalic' nucleus (symbolized as Vʔ) versus those not ending with glottal constriction, a non-glottalic nucleus (VV).<sup>17</sup>

If a stem rime lacked a coda consonant, the nucleus had to be full (VV or Vʔ),<sup>18</sup> but reduced as well as full vowel nuclei occurred in rimes ending in a consonant. A rime ending in

---

<sup>16</sup> Only stem rimes could consist of the segments from which tone evolved, so the extension of tone contrasts to prefixes is not discussed.

<sup>17</sup> Aside from tone, the reflexes of the glottalic and non-glottalic nuclei are identical, which indicates the both are full and that they differ only in that the second mora of the glottalic nuclei was produced with glottal constriction.

<sup>18</sup> Prefix rimes might consist of just a single reduced vowel nucleus.

a glottalic sonorant or fricative with a non-glottalic full vowel nucleus did not contrast with one with a glottalic full vowel nucleus. Otherwise, as Table 7 shows all possible combinations of rimes and nuclei occurred. The apostrophe marking a glottalic articulation is placed before the consonantal symbol for fricatives (’X) and sonorants (’R) but after that for stops K’) to indicate probable differences in the relative timing of the laryngeal and oral articulations between the first two manners of articulation and the third.

Nucleus	Stem-Final Consonant						
	None	Glottalic C			Non-Glottalic C’		
		Stop	Fricative	Sonorant	Stop	Fricative	Sonorant
VV	VV	VVK	VVX	VVR	VVK’	VV’X	VV’R
V?	V?	V?K	V?X= VV’X	V?R= VV’R	V?K’	V?X= VV’X	V?R = VV’R
V	–	VK	VX	VR	VK’	V’X	V’R

Table 7: Contrasting rime shapes in Proto-Athabaskan stems. “=” indicates that the flanking rime shapes are indistinguishable in their tonal behavior.

### 2.6.2 Tonogenesis

Table 8 shows how these Proto-Athabaskan rime properties interact to produce tones in two representative tonal languages, Chipewyan and Gwich’in, and how they are preserved in a non-tonal language, Hupa. Anticipating discussion to follow, the tone that developed in some stems ending in a glottalic consonant is referred to from now on as the “marked” tone; the other as the “unmarked” tone. This tone is high in Chipewyan but low in Gwich’in. The unmarked or opposite tone occurs in all other stems in both languages. The development of opposite tones from the same source is discussed after the properties of the rime that influence tonogenesis are accounted for.

The contrasts between Table 8(b,d) and (a,c) show that marked tone developed in reduced vowel stems ending in a glottalic stop or sonorant, while unmarked tone developed in those ending in a non-glottalic stop or sonorant. When the vowel was full, however, marked tone only developed in stems ending in glottalic sonorants, Table 8(i) versus (h) and not in those ending in glottalic stops, cf. Table 8(f,g) versus (e). In syllables with full vowels

ending in stops, unmarked tone developed regardless of whether the stop was glottalic, and the glottalic articulation was lost in the tonal languages. Table 8(k) versus (j) shows that marked tone also developed in stems whose nucleus was a glottalic full vowel. Table 8(l-n) show that marked tone developed uniformly in such stems regardless of whether the stem-final consonant was glottalic (m,n) or not (l).

Cognates in the non-tonal language Hupa to a large extent retain the original glottalic versus non-glottalic contrast in stem-final stops, which was lost in the reflexes of final stops in both tonal languages. Gwich'in retains the contrast in stems ending in sonorants (Table 8(d,i) versus (h,c)) and between glottalic and non-glottalic nuclei in stems ending in no consonant (Table 8(k) versus (j)). The loss of these latter contrasts in Chipewyan is in fact atypical; the majority of tonal Athabaskan languages retain both (see pp. 162-163 in Kingston, 2005, for examples). In syllables with these rimes, tone remains redundant on the glottalic contrast in the final sonorant or in the nucleus itself.

No glottalic fricatives are reconstructed for Proto-Athabaskan, but they are derived in certain stem allomorphs from glottalic stops. For example, the lengthened allomorph \*ʔaatʃ 'few go' becomes \*ʔaaʃ, which in turn appears with marked tone in Chipewyan ʔás and Gwich'in ʔòò, cf. the perfective allomorph, \*ʔaatʃ-ŋ, where the stem-final consonant remained unspirantized, and unmarked tone evolved in Chipewyan ʔáz and Gwich'in ʔóó. These derived glottalic fricatives behave like glottalic sonorants in producing marked tone on full as well as reduced vowel nuclei.

Kingston (1985, 1990, 2005) argued that sonorants and fricatives behaved differently from stops because they lacked a stop burst with which the laryngeal articulation could be coordinated. The laryngeal articulation in sonorants and fricatives could therefore shift more easily to the beginning of the consonantal articulation than in a stop, where it remains bound to the release at end of the closure. This timing difference would also explain why no contrast was possible in Proto-Athabaskan between VV and Vʔ nuclei in stems ending in glottalic sonorants or fricatives: the glottal constriction at the beginning of a sonorant or

Gloss	Proto-Athabaskan	Chipewyan High	Gwich'in Low	Hupa Non-tonal
Reduced vowel-Stop, VK versus VK' > unmarked versus marked tone				
a. 'smoke'	*ʔəd	ʔər	ʔád	ʔid
b. 'belly'	*wət'	bór	vàd	mət'
Reduced vowel-Sonorant, VR versus V'R > unmarked versus marked tone				
c. 'bone'	*ts'ən	tθ'ən	tθ'án	ts'iŋ
d. 'fire'	*qʷn'	kún	kʷ?	xoŋ'
Full vowel-Stop, VVK versus VVK' > unmarked tone				
e. 'wife'	*ʔaad	ʔà	ʔád	ʔad
f. 'scab'	*ʔuut'	ʔùr	ʔíd	ʔoh
g. 'spit'	*ʃ <sup>w</sup> eeq'	sèy	s'íg	see?
Full vowel-Sonorant, VVR versus VV'R > unmarked versus marked tone				
h. 'rain'	*k <sup>j</sup> aan	tʃà	tsín	k <sup>j</sup> aŋ
i. 'shit'	*k <sup>jw</sup> aan'	tsá	trìà?	tʃ <sup>w</sup> aŋ'
Full vowel-No final C, VV versus V? > unmarked versus marked tone				
j. 'stone'	*tsee	tθè	kíí	tsee
k. 'foot'	*qe?	ké	k <sup>w</sup> àj?	xe?
Glottalic full vowel-Stop, V?K and V?K' > marked tone				
l. 'flour'	*ʔeʔdʒ	ʔéz	ʔùh	ʔiim
m. 'vein'	*k <sup>j</sup> 'uʔts'	tʃ'ùdðé <sup>a</sup>	tʃ'iʔò	k <sup>j</sup> 'its'
n. 'pitch'	*dʒeʔq'	dzé	dzìh	dzeex

Table 8: Reflexes of the Proto-Athabaskan rime shapes in two tonal languages, Chipewyan and Gwich'in, and a non-tonal language, Hupa. <sup>a</sup> The low tone in the Chipewyan reflex of 'vein' is unexpected. The hook under a vowel indicates that it's nasalized.

fricative consonant would coincide with that at the end of a glottalic full vowel and render strings consisting of VV'R and VV'X indistinguishable from those consisting of V?R and V?X.

The coordination of the glottalic articulation with the release of a stop was not enough to prevent it from constricting all or most of a preceding reduced vowel, so marked tone developed in such stems, but in stems with full vowels, only a portion of the vowel would be constricted by coarticulation with the final consonant, too little, apparently, for marked tone to develop on such vowels.

### 2.6.3 Opposite tones from the same source

Kingston (2005) argued that opposite tones could have developed from stem-final glottalic consonants in Chipewyan and the other high-marked Athabaskan languages versus Gwich'in and the other low-marked languages because that those consonants' laryngeal articulation could have been pronounced so as to raise or lower F0. If the glottal constriction was achieved by only contracting the thyroarytenoid muscle, then the voice quality of adjacent vowels would be creaky and F0 would be characteristically low; however, if the cricothyroid muscle was also contracted at the same time, the adjacent vowel's voice quality would instead be tense and its F0 would be characteristically high. Although thyroarytenoid contraction ordinarily raises F0, it would not do so in either pronunciation of glottal constriction because this muscle is contracted so strongly to achieve sufficient medial compression of the vocal folds that vibration is limited to the folds' covers and does not extend deeper, into their bodies. Cricothyroid contraction thus determines whether the covers are stretched and F0 is raised or not. The choice between the two alternative pronunciations in a particular Athabaskan language was apparently arbitrary and idiosyncratic because whether a language is high- or low-marked does not correspond to any other sound changes used in sub-grouping the family.

According Leer (1999), the high- and low-marked Athabaskan languages of Canada and Alaska can be separated geographically, with the high-marked languages on the east side of the Canadian Cordillera, extending into the Mackenzie River drainage and the low-marked languages on the west side of the Cordillera up into the Yukon and Tanana River drainages to the north and south into British Columbia. Yet within each of these two groups are found closely related languages with the opposite tone, e.g. low-marked Dogrib is spoken among otherwise high-marked and closely related languages in the Mackenzie River drainage, while in the low-marked Yukon and Tanana River drainages, Northern Tutchone is high-marked next to its close relative, low-marked Southern Tutchone and similarly Tanacross is high-marked next to its close relatives, low-marked Upper and Lower Tanana.

The glottalic versus non-glottalic contrast must have been lost very early in stem-final

stops in the tonal languages as none of them retain any vestige of it in the consonants themselves. How then could these recent reversals come about, to produce low-marked Dogrib and high-marked Northern Tutchone and Tanacross?

The answer probably lies in the retention of the contrast in stem-final sonorants as well as between glottalic and non-glottalic full vowels in these languages down to the present day. If the ancestor of the tonal languages originally split into a high-marked and a low-marked dialect when some of its speakers chose to pronounce all the glottalic consonants with cricothyroid as well as thyroarytenoid contraction while others chose to do so with thyroarytenoid contraction alone, then there is no reason why speakers should lose this freedom to choose how to pronounce glottalic consonants in the subsequent history of the family. Speakers of Dogrib, Northern Tutchone, and Tanacross could have exercised this choice quite recently in words ending in glottalic sonorants or whose rime consisted of a glottal nucleus, where tone remained redundant on the laryngeal articulation of the stem-final consonant or the nucleus. But what of stems that ended in stops in Proto-Athabaskan, where the glottalic versus non-glottalic contrast was lost long before Dogrib, Northern Tutchone, or Tanacross diverged tonally from their near relatives? The contrastive tones in what were once stop-final stems must have switched value at the same time as those in the stems where the tones were predictable. That is, all high and low tones were treated alike regardless of whether they were contrastive or redundant.

The parallel behavior of the two kinds of tones is likewise observed in the languages that have not recently reversed tone values, where no speakers have exercised their freedom to choose how to pronounce the glottalic sonorants or nuclei and only reversed tone values in stems whose rimes contain these segments. This parallelism indicates that the phonologization of the sound change as high or low tones in the stems that ended in stops was extended to the pronunciation of the tones that remained predictable in stems ending in glottalic sonorants or whose rime consisted of a glottalic nucleus. Phonologization also ensured that any change in the pronunciation of the tones in the latter kind of stems was therefore reciprocated

by an identical change of the phonologically identical tones in the stems that had originally ended in stops. The persistence of this linkage between the sound change's phonetic motivation and its reflexes' subsequent behavior undermines Blevins's (2004) claim that sound changes are severed from their phonetic motivations once they have been phonologized.

## 2.7 Cross-linguistic variability in tonogenesis

Athabaskan is by no means unique in producing the opposite tone values from the ostensibly the same source. Both [ʔ] and [h] (or [constricted glottis] and [spread glottis]) have produced both high and low tones on preceding vowels.

A rising tone arose in syllables ending in a glottal stop in Vietnamese (Table 6, but in Lhasa Tibetan, a falling tone has emerged in such syllables instead (Mazaudon, 1977). Mazaudon proposes that a rising or high tone emerges before a true [ʔ], in which the glottis is tightly closed, while a falling or low tone arises when the glottal stop is instead realized with incomplete glottal closure, i.e. as creaky voice. The discussion of Athabaskan earlier in this section indicated that whether a higher or lower tone evolves is not a matter of whether the glottis is completely closed but instead whether the cricothyroid as well as the thyroarytenoid muscle is contracted when the glottis is constricted. Mazaudon's suggestion nonetheless resembles Kingston's in attributing the difference in which tone emerges to phonetic differences in the pronunciation of the source consonant or its laryngeal articulation.

Regarding [h], recall that a falling tone emerged in Vietnamese in syllables that ended in [h] in Proto-Vietic, while a high tone emerged in Utsat in syllables that ended in [h] in Proto-Cham. A high tone also emerged after voiceless aspirated stops, [s], and [h] in Korean, and on Punjabi vowels followed by [h] or breathy voiced stops in Middle Indic (Ohala, 1973). Lower F<sub>0</sub> is usually associated with breathy voice, as in tonal W. Kammu compared to registral W. Kammu or Lamet, and in E. Cham compared to W. Cham. Moreover, obstruents that are often described as breathy voiced “depress” tone in many S. Bantu languages,<sup>19</sup> where

---

<sup>19</sup> In many instances, the depressors are no longer particularly breathy, nor even voiced (Schachter, 1976; Traill, 1990; Jessen & Roux, 2002; Maddieson, 2003; Strazny, 2003; Downing & Gick, 2005; Downing &

following high tones are realized as rising tones, are displaced onto following syllables with a low tone appearing in their place, and high tones fail to spread or displace across breathy voiced consonants (Rycroft, 1980, 1983; Traill et al., 1987; Traill, 1990; Cassimjee, 1998; Cassimjee & Kisseberth, 1998; Donnelly, 2009). These observations suggest that the /h/ which lowers F0 of adjacent vowels may be voiced [fi], which is actually breathy voiced, while that which raises their F0 is instead voiceless [h].

Alternatively, a higher F0 could be produced at the same time as the glottis is abducted by contracting either the cricothyroid or the thyroarytenoid. Cricothyroid contraction would raise F0 by stretching the vocal folds' covers, as it does in the production of tense voice. Thyroarytenoid contraction stiffens the vocal fold body and would raise F0 if vibration extended deep enough into the body. Thyroarytenoid contraction is also used to press the ligamental portions of the folds medially against one another in whisper and whispery voice, where the cartilagenous portions of the vocal folds are abducted by posterior cricoarytenoid contraction and interarytenoid and lateral cricoarytenoid relaxation (Laver, 1980). Whisper is produced when the medial compression is great enough to prevent any vibration of the adducted ligamental portion of the folds, and whispery voice when it's not so extreme. Either voice quality would elevate F0 in adjacent vowels.

## 2.8 Interim discussion

In this section, I have described how tone emerged from the laryngeal articulations of following consonants in the earliest ancestors of present-day Vietnamese, the Chinese languages, and the tonal Athabaskan languages – [constricted glottis] in all three families as well as [spread glottis] in the ancestors of Vietnamese and Chinese. In more recent ancestors of present-day Vietnamese and Chinese, the tones that arose from this source subsequently split under the influence of [voice] preceding consonants. Tonogenesis in Athabaskan shows

---

Schadeberg, 2009). As a result, depression has become a nearly or completely autonomous feature from the original laryngeal articulation that first produced it (Traill, Khumalo, & Fridjhon, 1987; Traill, 1990; Maddieson, 2003), much like tone itself has in Vietnamese or Chinese.

that a glottal constriction can be pronounced so as to either lower or raise F0, depending on whether the cricothyroid as well as the thyroarytenoid is contracted. Other examples cited above show that constricting the glottis can either raise or lower F0 in other languages, too, as can spreading the glottis. In the next section, we will see that such cross-linguistic variability is pervasive. This variability is another, strong hint that the so-called “perturbations” of F0 caused by these various laryngeal articulations may instead be controlled. Unlike the F0 lowering that typically accompanies voicing in obstruents, this manipulation of F0 does not apparently integrate perceptually in any obvious way with other acoustic correlates of these contrasts, but instead appears only to add another phonetic difference to consonants specified for [constricted glottis] or [spread glottis].

### 3 Tone splitting in East and Southeast Asia

#### 3.1 Two- and three-way splits and cross-linguistic variability

Tones first appeared in the ancestor of modern Vietnamese under the influence of the laryngeal articulations of coda consonants, if any, and then the three tones that arose from this source split in two under the influence of the laryngeal articulations of onset consonants, ultimately producing six contrasting tones in the present-day language (at least as it’s spoken in the northern part of Vietnam). The first step is tonogenesis proper, while the second is tone splitting. Tone splitting and the resulting multiplication of tones under the influence of onset consonants is extremely common among the languages of East and Southeast Asia, where it has occurred in the histories of languages in the Sino-Tibetan, Hmong, Tai, and Kam-Sui families, as well as the examples already discussed above.<sup>20</sup> Both two- and three-way splits occur, where in two-way splits (Table 9), [voice] onset consonants induce distinct tones from consonants with all other laryngeal articulations, and in three-way splits (Table 10), [spread

---

<sup>20</sup> The examples and analyses presented in this section reflect a part of the work that David Solnit and I did together. We could not have accomplished even a small fraction of what we did without the exceptionally clear and insightful accounts of tonogenesis in East and Southeast Asian languages that can be found in Haudricourt (1972); Mazaudon (1977); Matisoff (1972, 1973).

glottis] and [constricted glottis] as well as [voice] consonants each induce distinct tones.

Languages differ in whether the tonal reflexes obtained after [voice] consonants are lower or higher than those obtained after consonants with other laryngeal articulations (Table 9 versus Table 11), and they likewise differ in whether the tonal reflexes obtained after [spread glottis] consonants are lower or higher than those obtained after [constricted glottis] consonants (Table 10(a) versus (b)).<sup>21</sup> Before providing specific examples, it's worth emphasizing that the examples presented are simply a few of many, many possible examples of these developments, so none of these developments can be dismissed as rare or exceptional.

Among the three Hmong languages illustrated in Table 9, Shu wei's tones have not split (Table 9(a)), while Shih men k'an (Table 9(b)) is at the first stage in the split, in that distinct reflexes have evolved after [voice] consonants than after consonants with other phonation types, but these tones all remain predictable allotones because none of the [voice] consonants have merged with consonants of other phonation types. Finally, in Chi wei (Table 9(c)), some of the tonal reflexes are now contrastive, because \*b has merged with \*p before reflexes of proto-tone A and \*ʔm with \*m before reflexes of all proto-tones.<sup>22</sup>

In Yung-chiang Kam and Nakhorn Sithammarat Thai (Haudricourt, 1972), lower tones have evolved in syllables that began with voiced stops and sonorants than other consonants, but the two languages differ in whether higher tones have evolved after [constricted glottis] consonants – Yung-chiang Kam (Table 10(a)) – or [spread glottis] consonants – Nakhorn Sithammarat Thai (Table 10(b)). The voiceless unaspirated stops pattern with the [constricted glottis] consonants in both languages – in other languages, they pattern with the [spread glottis] consonants instead. The patterns of neutralization of the original phonation contrasts are both similar and different in the two languages: in both, [spread glottis]

---

<sup>21</sup> Proto tones are identified with the conventional letters, A-D. For Chinese languages they correspond to *ping*, *shang*, *qu*, and *ru* tones. For Tai and Kam-Sui, they are used as in Li (1977) and Li (1948, 1965), respectively. For Lolo-Burmese, A-C correspond to tones 1-3 in Matisoff (1972, 1979) and Bradshaw (1979).

<sup>22</sup> In these and other examples presented in this section, the same tone may appear as the reflex of more than one proto-tone, sometimes in more than one consonantal context. I assume here without argument or explanation that these identities reflect later mergers.

Low		≠	High		
voice	b, m, v	≠	p		
			p <sup>h</sup> , h <sub>m</sub> , f	spread	
			?b, ?m	constricted	
Proto tones					
		A	B	C	D
(a) Shui wei					
no split		31	55	35	31
(b) Shih men k'an*					
High: (m)p, (m)p <sup>h</sup> , h <sub>m</sub> , ?m		55	55	33	11
Low: b > b/b <sup>h</sup> , v > v/v <sup>h</sup> , mb, m		35	35/13	53/31	31
(c) Chi wei**					
High: (m)p, (m)p <sup>h</sup> , h <sub>m</sub> , ?m > m		35	44	53	44
Low: b > p/p <sup>h</sup> , mb > m/ <sup>h</sup> m, m > m/ <sup>h</sup> m		21	22	42	22

Table 9: Two-way tone splitting in three Hmong languages: (a) Shui wei, no split, (b-c) Shih men k'an and Chi wei, voiced-low splits. \* Voiced stops become breathy voiced in Shih men k'an before reflexes of proto tones A and D but retain modal voicing before B and C reflexes; the higher variant of the low reflexes of tones B and C occurs in nouns, the lower in verbs. \*\* In Chi wei, original voiced stops simply devoice in proto tone A but acquire breathy voicing before reflexes of proto tones B, C, and D. 'x > y' indicates that 'x' changed into 'y' in the course of the split. Symbols for labial consonants are used in this table and others below to represent the various phonation types involved in the splits.

sonorants merge with [voice] sonorants; in Yung-chiang Kam, the [voice] stops merge with the voiceless unaspirated stops, while in Nakhorn Sithammarat Thai, they instead merge with the voiceless aspirated stops; and in Yung-chiang Kam but not Nakhorn Sithammarat Thai, the [constricted glottis] stops have merged with [voice] sonorants. For brevity, the outcomes illustrated by these two languages are referred to henceforth as “Constricted High” and “Spread High” splits.

In all the examples thus far, lower tones have consistently evolved after [voice] consonants that those of other phonation types, but as the examples in Table 11 show, some languages exhibit higher reflexes instead, both when tones split two ways (Table 11(a)) and three ways (Table 11(b)). These outcomes can be referred to as “Voiced High” splits, as compared to the “Voiced Low” splits in Tables 9 and 10. In these cases, too, whether an original phonation

		Proto tones		
		A	B	C
b, m, v	≠ p <sup>h</sup> , h <sub>m</sub> , f, (p)	≠	ʔb, ʔm, (p)	
voice	≠ spread	≠	constricted	
<hr/>				
(a) Constricted High				
Yung-chiang Kam				
High:	p, ʔb > m, ʔm > m	55	53	323
Mid:	p <sup>h</sup> , h <sub>m</sub> > m	35	453	13
Low:	b > p, m	212	33	31
<hr/>				
(b) Spread High				
Nakhorn Sithammarat Thai				
High:	p <sup>h</sup> , h <sub>m</sub> > m	53	53	55
Mid:	ʔb, p	24	24	33
Low:	b > p <sup>h</sup> , m	31	35	11

Table 10: Three-way tone splitting in a language where [constricted glottis] consonants induce higher reflexes than [spread glottis] consonants, Yung-chiang Kam (AKA “Tung”), and in one where the [spread glottis] consonants induce higher reflexes, Nakhorn Sithammarat Thai.

type merges with another appears to be unrelated to the levels of the tones that evolve from those phonation types.

(a) Two-way					
Shan					
		A	B	C	
High:	b > p, m	55	22	44	
Low:	p, h <sub>m</sub> > m, ʔb > m/w	334	11	22	
<hr/>					
(b) Three-way					
Szu ta chai					
		A	B	C	D
High:	b > p, m	35	33	44	33
Mid:	(m)p, ʔm > m	33	44	13	31
Low:	p <sup>h</sup> , mp <sup>h</sup> > mp <sup>fi</sup> , s > z <sup>fi</sup> , h <sub>m</sub> > fi <sub>m</sub>	13	34	11	11

Table 11: Voiced High two-way and three splits in Shan and Szu ta chai.

In the face of such dramatic crosslinguistic variation, it’s perhaps easy to ignore the three outcomes that are consistent across all these examples (as well as the many others that could be presented):

1. Distinct reflexes emerged after original [voice] consonants than consonants of other

specified phonation types,

2. Distinct reflexes emerged after the one series of consonants that is unspecified for phonation type, the voiceless unaspirated stops, than the [voice] consonants,
3. Sonorants and obstruents of a given phonation type produce the same reflexes.

Aside from whether tones split in two or three, languages do not differ in how their present-day tonal reflexes can be grouped by the classification of consonants by phonation type or laryngeal articulation in their respective proto-languages. What differs between languages is instead the **level** of the tone that emerged after a consonant of a particular phonation type or laryngeal articulation.

### 3.2 An explanation

There are two essentially similar phonetic puzzles here. First, how can higher reflexes evolve in some languages after [voice] consonants than consonants with other phonation types when [voice] *obstruents* lower the F0 of following vowels? Second, how can higher reflexes evolve after [constricted glottis] than [spread glottis] consonants in some languages while in others higher reflexes evolve after [spread glottis] than [constricted glottis] consonants? The traditional answer to the first question is that the present-day Voiced High languages were originally Voiced Low, in conformity with phonetic expectations, but after the initial [voice] contrast neutralized and the tone split became contrastive, the tones' values were no longer constrained phonetically, and the tones that had evolved after the original [voice] consonants exchanged values with those that evolved after other consonants. A similar story might be concocted for the Spread High versus Constricted High difference: either [spread glottis] or [constricted glottis] consonants first induced higher tone reflexes in the three-way splits, and the higher reflexes later exchanged values with the lower reflexes.

Kingston & Solnit (1989) give three arguments for ruling out such post-merger exchange processes. First, exchange processes otherwise do not occur but would have to apply fre-

quently in this linguistic area to this linguistic property.<sup>23</sup> Second, since it is the tone features that have developed on the vowels that invert their values, all proto-tones would have to undergo the process. Once inversion occurred, the prior, uninverted stage would no longer be detectable, and there would be no way to know it had ever happened. If the mechanism that inverted tones could apply to individual proto-tones, inversion would be observable in that some proto-tones would exhibit post-inversion Voiced High reflexes, while others would retain the prior Voiced Low reflexes. However, such mixed languages do not occur (see also Brown, 1975, for further elaboration of this argument). Third, we can observe the difference between Spread High and Constricted High splits prior to the merger of the laryngeal contrasts in the consonants. In the Spread High language Nakhorn Sithammart Thai (Table 10b), the [spread glottis] and [constricted glottis] consonants still contrast, as they do in the Constricted High Karenic language, Pwo (Table 12). If languages can differ in whether [spread glottis] or [constricted glottis] consonants induce higher tone reflexes *before* the laryngeal contrasts merge, i.e. while the tone reflexes still remain predictable from the consonants' laryngeal articulations, then there is no need to assume that one development is original and the other a later inversion. A Voiced Low split without loss of the [voice] contrast can be observed in Shih men k'an Hmong (Table 9b above), but I have not found any cases of a Voiced High split in which the [voice] contrast has not merged.

	A	B
ʔb, p	33	41
p <sup>h</sup> , h <sub>m</sub> > m	11	41
b > p <sup>h</sup> , m	33	24

Table 12: Three-way Constricted High tone split in Pwo.

A further observation led Solnit & Kingston (1988) and Kingston & Solnit (1989) to propose an alternative to exchanges, namely, none of the physiological mechanisms that

---

<sup>23</sup> A related argument is that exchanges are too powerful to be countenanced and cannot be implemented in optimality theory (Alderete, 2001; Moreton, 2004): if a constraint ranking in a language will cause A to be replaced by B, there cannot also be a ranking in that language that will cause B to be replaced by A.

might lower F0 automatically as an accidental consequence of an articulation intended to sustain voicing in a [voice] obstruent are needed to produce (modal) voicing in sonorants, because air flow out of the oral cavity is not impeded and  $P_o$  does not build up. Yet, as noted above, voiced sonorants always pattern with the corresponding [voice] obstruents in tone splits. This uniformity suggests that voiced sonorants are also specified [voice], or that what matters in tone splits is that they are phonetically voiced. As sonorants of all other phonation types also pattern together with the corresponding obstruents in tone splits (the third observation above), and there is likewise no more reason why spreading or constricting the glottis in their production should necessarily raise or lower F0 than in [h] or [ʔ] (but cf. Maddieson, 1984), then it is possible that the sonorants of a particular phonation type could have been pronounced differently in proto-languages from which higher tones developed than in those from which lower ones did – the same argument was made above for the development of opposite tones from the same phonological sources in different Athabaskan languages.

Once the pronunciation of the sonorants and its effect on F0 had been determined in these languages, the corresponding obstruents would be pronounced and affect F0 in the same way, even if that pronunciation would conflict with the obstruent's phonetic proclivity to perturb F0 in the opposite direction. A phonetic conflict may in fact only arise with [voice] obstruents, which lower F0, as both [spread glottis] and [constricted glottis] obstruents can be pronounced so as to raise or lower F0.

Solnit & Kingston and Kingston & Solnit formalized this hypothesis by specifying consonants for tones, i.e. for [stiff] and [slack] as well as for [voice], [spread glottis], or [constricted glottis]. The highest reflexes would develop after the consonants specified as [stiff], the lowest after those specified [slack], and the intermediate reflexes after consonants not specified for either glottal tension feature. Although this formalization used the features introduced by Halle & Stevens (1971), [stiff] and [slack] were not substituted for [voice]. Instead, a consonant's value for [stiff] and/or [slack] was acknowledged to be predictable from/redundant on its value for [voice], [spread glottis], or [constricted glottis] within a language, but not

cross-linguistically.

Two important facts are not adequately accounted for by this proposal: entire language families or substantial sub-groups within them in East and Southeast Asia are either Constricted High or Spread High, but Voiced Low and Voiced High developments are found within the same families, sub-groups, and even between closely related languages. In other words, it appears that speakers arbitrarily chose to produce higher F0 values after [constricted glottis] compared to [spread glottis] consonants or vice versa just once in the history of a family or sub-group and later speakers did not deviate from that choice, while speakers have repeatedly chosen to produce higher or lower F0 after voiced consonants throughout the history of these languages.

This difference is surprising for two reasons. First, only voiced obstruents are expected to have a consistent and automatic lowering effect on F0, yet they appear to be most labile in tone splits. Second, both two- and three-way splits set voiced consonants apart from consonants with another phonation type, but only three-way splits distinguish between constricted and spread glottis consonants. Yet the voiced consonants can produce reflexes that are both higher and lower than those produced by consonants of the other two phonation types.

Sonorants behave differently in these East and Southeast Asian tone splits than in Athabaskan tonogenesis. On the one hand, in East and Southeast Asia, transfer of contrast from laryngeal articulations in consonants to tone in vowels is often complete for syllables beginning with sonorants but remains incomplete for syllables beginning with obstruents, where laryngeal contrasts are retained and the tone split on the following vowels remains redundant. In Athabaskan, on the other hand, it's the sonorants that retain their laryngeal contrast and the obstruents that have transferred the contrast from their laryngeal articulations to tone on the preceding vowel.

### 3.3 Tones from preceding versus following consonants: Splits versus tonogenesis

Tonogenesis from laryngeal contrasts in preceding consonants in languages that previously did not contrast syllables for tone, such as has occurred in W. Kammu dialects, E. Cham, and Utsat, does not in fact occur in Vietic, Sino-Tibetan, Hmong, Tai, or Kam-Sui; only tone splitting did. With the exception of these three languages and Yabem,<sup>24</sup> preceding consonants are apparently not capable of creating tones where there were none before, but only of multiplying existing tones.

There is actually a paradox here. Hombert (1977) showed that the F0 perturbations induced by an initial [voice] contrast were much smaller and shorter lasting in Yoruba, where high, mid, and low tones contrast, than in English, where no contrasts between morphemes are carried by tone. The perturbations may be constrained in size and extent in Yoruba to prevent their so altering F0 in the following vowel that one tone would be confused with another. If the perturbations induced by initial consonants' laryngeal articulations are generally constrained in tone languages, then how could they ever bring about the splits just described? The answer to this question must be that speakers of these languages had already been sensitized to F0 differences by the tones that arose earlier from final consonants' laryngeal articulations. Once they were paying attention to such differences, they would be more likely to attend to and perhaps rely on the systematic F0 differences after preceding consonants, too, even if those differences were at first constrained in size and extent. This shift in attention led to the F0 differences supplementing and eventually replacing the original laryngeal contrasts in the preceding consonants.

This line of reasoning raises another paradox: how then could tone have developed from an initial [voice] contrast in the W. Kammu dialects, E. Cham, or Utsat, where there was no tone before? Tone could not have first developed from the laryngeal articulations of final

---

<sup>24</sup> Given that accentual phrases in Korean are characterized by a THL melody, this language might already have been sufficiently tonal for the differences in the laryngeal articulations of initial consonants to determine whether the initial tone in this melody is high or low.

consonants in any of these languages, because the influence of such consonants on the F0 of preceding vowels remains entirely allophonic in all but one instance, the Utsat high tone. The resolution of this paradox has two parts. First, the F0 perturbations of the following vowel induced by the initial [voice] contrast could have been larger and lasted longer precisely because the ancestors of these languages were not tonal. Second, all three languages have been in extensive and sustained contact with other languages whose tones have split under the influence of an initial [voice] contrast, and this contact could have led their speakers, too, to reinterpret the F0 perturbations associated with an initial [voice] contrast as a tone contrast instead. The development of tone in 6th century Vietnamese has also been argued to be a product of contact with a tone language, in its case, Middle Chinese (Haudricourt, 1954b).

There are, of course, two other languages where the F0 perturbations induced by initial laryngeal articulations have been sufficient to create tone *de novo* on following vowels, Yabem and Korean. Both developments are less likely to have been spurred by language contact than in the languages just discussed. Yabem does not appear to be in contact with any other tonal languages, aside from its close relative, Bukuwa. Tone developed in the same way in Bukuwa as Yabem, presumably before they became distinct languages, but since their separation, stops have devoiced in Bukuwa in the onsets of strong syllables in feet, and voiced in the onsets of weak syllables. Tone and onset voicing no longer covary in Bukuwa as they still do in Yabem. Korean has, of course, been in constant and sustained contact with Chinese, but the phonologization of the F0 differences after [spread glottis] and [constricted glottis] versus unspecified consonants is too recent to be explained by that contact. In both Yabem and Korean, however, the tonal differences remain redundant on the initial consonants' laryngeal articulations. The redundancy is complete in Korean, though the tone difference may soon be contrastive if the reduction in VOT differences between aspirated and lax stops reported by Silva (2006) continues. These languages then are genuine or at least incipient exceptions to the generalization that contrastive tone does not develop from laryngeal articulations

(perhaps more narrowly [voice] contrasts) in preceding consonants, although it has done so very frequently from the laryngeal articulations of following consonants.

## 4 Tones from uncommon sources

While tones have frequently developed or split in vowels under the influence of the laryngeal articulations of neighboring consonants, only two languages appear to have split under the influence of the vowels' own articulations. All things being equal, F0 is higher in vowels in which the tongue is higher than in vowels in which the tongue is lower. This variation in the "intrinsic" F0 of vowels is so pervasive that it is thought to a universal, mechanical consequence of differences in tongue height, where raising the tongue pulls the larynx up via the hyoid bone (Ohala & Eukel, 1987; Whalen & Levitt, 1995; Whalen, Gick, Kumada, & Honda, 1999), but see Ladd & Silverman (1984); Steele (1986); Kingston (1991, 1992) for contrary evidence. Given these intrinsic F0 differences' ubiquity and apparently automatic character, it would not be at all surprising if tones developed from them. Yet it appears that, with just two exceptions, tone never has developed from this source – other putative cases have been argued not to be products of intrinsic F0 differences between vowels differing in tongue height (Hombert, 1978; Maddieson, 1978; Schuh, 1978). The exceptions are the Angkuic language, U, a not-too-distant relative of the Kammu languages discussed in §2.3.1 above (Svantesson, 1989), and Lugbara, a Moru-Madi language (Central Sudanic branch of Nilo-Saharan), spoken in the Congo-Uganda-Sudan border area (Andersen, 1986).

### 4.1 Tone from vowel height: U

In U, the high level tone in the closely related languages, Hu and Lamet, has split: it remains high in syllables with high vowels, but became low in syllables with non-high vowels (Table 13a-d versus e-g, Svantesson, 1989).

### 4.2 Tone from advanced tongue root: Lugbara

Western Lugbara (Andersen, 1986) distinguishes four tones: low [fɪ] 'it exploded', mid [fɪ] 'he entered', high [fɪ] 'intestines', and extra-high [fɪ] 'they entered', and seven vowels /i, ɪ, ɛ,

gloss	U	Hu	Lamet
a. louse	nchí	nsíʔ	síʔ
b. rope	sí	(pa)síʔ	(pl)síʔ
c. hole	nthú		ntuʔ
d. vegetable	thú		tɥuʔ
e. tree	sì	θéʔ	khɛʔ
f. rice	ŋkhù	ŋkhóʔ	
g. yesterday	khù	sŋkhóʔ	

Table 13: High tone in U where the vowel is high in Hu and/or Lamet (a-d) versus low tone when the vowel is non-high in the related languages (e-g).

a, ɔ, ʊ, u/. The vowels /ɛ, ɔ/ are raised to [e, o] when followed by /i/ or /u/ within a word, and /i, u/ do not co-occur with /ɪ, ʊ/ within a word. Andersen proposes that the language therefore distinguishes two sets of vowels on the surface, which differ in whether the tongue root is advanced [ATR]: [+ATR] = [i, e, a, o, u] versus [-ATR] = [ɪ, ɛ, a, ɔ, ʊ].

In the second syllable of disyllabic verb stems in this language, both [-ATR] and [+ATR] vowels can bear low or mid tone, but high and extra-high tone are in complementary distribution: high tone occurs when the vowel is [-ATR], [áví] ‘play’ versus extra-high tone when it’s instead [+ATR] [átsí] ‘walk’. In nouns, the distribution of low and mid tones is also unconstrained by the vowel’s ATR specification, but high tone does not occur on [+ATR] vowels and the extra-high tone can occur occur on [-ATR] as well as [+ATR] vowels, e.g. high [àrí] ‘blood’ but \*[àrí] and [òtsé] ‘dog’ versus [òdú] ‘thigh’. Andersen proposes that the high tone split in Western Lugbara and became high when the vowel bearing it was [-ATR] but became extra-high when that vowel was [+ATR]. Although the ATR contrast in mid vowels maintained by related languages has merged in W. Lugbara, correspondences between the ATR value of mid vowels in related languages and the distribution of extra-high versus high tone in Western Lugbara are too irregular to treat the tone split as compensating for this merger. The most one can say is that the distribution of these two tones can be predicted from the [ATR] values of vowels in present-day Lugbara.

Higher F0 has been associated with [+ATR] in other languages (Denning, 1989), so ATR

differences may affect vowels' intrinsic F0 much like vowel height does. Perhaps, it does so because [+ATR] vowels are often also somewhat higher than [-ATR] vowels.

### 4.3 Discussion

Tone splits from vowel height or ATR contrasts are decidedly rare. U and Lugbara are the only two languages known where they have occurred. Why should a vowel's own intrinsic F0 be so unable to split its tone, while F0 perturbations brought about by neighboring consonants' laryngeal articulations can so readily do so? The most plausible answer to this question is that intrinsic F0 differences between vowels differing in height depend on the vowels being prominent (Reinholt Peterson, 1978; Ladd & Silverman, 1984; Steele, 1986) or bearing a high(er) rather than low(er) tone (Hombert, 1977; Zee, 1980; Connell, 2002), while the effects of consonants' laryngeal articulations on the F0 of adjacent vowels do not appear to depend on the vowels' prominence or tones. If the intrinsic F0 differences between vowels are less consistent across prosodic contexts than the F0 differences between consonants, learners would be less likely to reinterpret them as differences in tone.<sup>25</sup>

## 5 From stress and intonation to tone: Tonogenesis in Germanic

Before concluding this chapter, I sketch the probable paths through which tone evolved in three North Germanic languages, Swedish, Norwegian, and Danish (Gårding, 1977; Riad, 1998; Lahiri, Riad, & Jacobs, 1999; Riad, 2000, 2003), and in Central Franconian (Hermans, 1999; Gussenhoven, 2000). Unlike all the other examples discussed in this chapter, the sources of tone in these languages are not the laryngeal articulations of segments, although the distribution of the tones is at least partly regulated by or correlated with various properties of the segments that make up the syllables on which the tones are expressed. The lexical tones in these languages instead come from the F0 correlate of stress in the Scandinavian languages and from intonation in Central Franconian.

---

<sup>25</sup> We cannot appeal here to the bias toward learning modular constraints proposed by Moreton (2008a,b, to appear) because the constraints referring to tone and consonants' laryngeal articulations are no less a-modular than those referring to tone and vowels' height.

## 5.1 North Germanic word accents

Words in both Swedish and Norwegian contrast in the tone contour expressed across them. These contours in both languages are best described as “word accents” because they characterize entire words, and they are aligned with the words’ stressed syllables. Following common usage, I refer to them as “Accent 1” versus “Accent 2”.

I use the facts of the conservative variety of Swedish spoken in and around Stockholm, i.e. Central Standard Swedish, to exemplify these word accents. Other dialects of Swedish, as well as Norwegian dialects, differ in whether the accents still contrast and the timing of the F<sub>0</sub> events that realize the accents relative to the segment strings of the words bearing them (Bruce & Gårding, 1978; Bruce, 1999; Gussenhoven & Bruce, 1999). Despite these differences between dialects and languages, their various synchronic word accent patterns all descend from the same historical source, the reinterpretation of the F<sub>0</sub> correlates of stress as word accents in late Proto-Nordic between 800-1200 CE.

On the surface, it appears that both word accents in Central Standard Swedish can be analyzed as a sequence of a H and L tone (Gussenhoven & Bruce, 1999, among many others). In words bearing Accent 1, the L tone is aligned with the stressed syllable, while in words bearing Accent 2, it is the H tone that is aligned with that syllable. Thus, the contrast between the two word accents is apparently not in the tones of which they are composed but the alignment of those tones with respect to the segments making up the words bearing them or more precisely that word’s primary stressed syllable (Bruce, 1977). The phonetic consequence of this difference in alignment is that F<sub>0</sub> is low from the beginning of the stressed vowel in words bearing Accent 1 (1a), while it starts high and falls to a low value across that syllable in words bearing Accent 2 (1b). For words bearing Accent 1, the H only appears if there is a preceding unstressed syllable, such as *med* ‘with’ in (1a).

- (a)  $\begin{array}{c} \text{H} \\ | \\ \text{m e d} \# \# \text{a n d e n} \end{array}$       ‘with the duck’ Accent 1      (1)
- (b)  $\begin{array}{c} \text{H L} \\ | | \\ \text{m e d} \# \# \text{a n d e n} \end{array}$       ‘with the ghost’ Accent 2

In both Accent 1 and Accent 2 words, a H tone appears immediately after the word accent tones on the last word of a focused constituent, and initial and final L% boundary tones appear at the beginning and end of each intonational phrase. The final L% tone is always realized as a fall in F0 at the end of the phrase, but the initial L% boundary tone may not be realized in an phrase containing an Accent 1 word if no more than a single unstressed syllable precedes the stressed syllable. As a result, the contour for (1a) starts with the high F0 of the word accent H on *med*, falls to the low F0 of the word accent L at the beginning of *an*, rises to the high F0 of the focal H at the beginning of *den* and finally falls again to the low F0 of the final L% boundary tone by the end of that syllable. The tonal composition of the intonation phrase in (1b) is identical, and the F0 contour that realizes those tones is nearly the same, except that all F0 targets are realized later. The contour begins with a rise from the low F0 of the initial L% boundary tone on *med* to the high F0 of the word accent H at the beginning of *an*, falls to the low F0 of the word accent L at the end of *an*, rises once more to the high F0 of the focal H on *den*, and falls at last to the low F0 of the final L% boundary tone. The initial L% boundary tone can be realized at the beginning of an unstressed syllable preceding an Accent 2 word because its word accent tones begin later than those in an Accent 1 word. Their later realization also delays the final rise and fall of the focal H and final L% boundary tones in Accent 2 compared to Accent 1 words.

The phonological analysis of these facts proposed by Riad (1998, 2003) differs substantially from this superficial account. Rather than a contrast in the tune-to-text alignment of a HL contour, words contrast privatively for the presence or absence of a lexical H\* tone. Words with Accent 1 have no lexical H\*, while words with Accent 2 have a specified H\* aligned with their primary stressed syllable. Focus is marked by a H tone, which follows the

lexical H\*, if any. The L that invariably precedes this H is not lexical but instead either part of focus marking, a default L, or one inserted to separate the lexical and focal Hs in Accent 2 words and satisfy the OCP. Riad describes intonational phrases containing Accent 1 words as having just a single F0 peak, that of the focal H, and therefore does not provide an account of the initial high F0 on an unstressed syllable preceding the stressed syllable of a word with this accent. The melody of an Accent 1 word is [LH]<sub>Focus</sub>-L%, while that of an Accent 2 word is H\*-[LH]<sub>Focus</sub>-L%.

Prior to the development of word accents in Proto-Nordic, words underwent a number of reducing sound changes. Of particular importance was syncope of vowels in light (CV) syllables (6th-9th centuries), e.g. heavy-light-heavy *'doomi,jan* > heavy-heavy *\*'døø,man* 'to judge' and heavy-light *'gastiz* > *'gæstr* 'guest'. As indicated, stress was quantity-sensitive at this stage, and any heavy syllable would have been stressed. Syncope of medial vowels created stress clashes whenever the syllables that became adjacent as a result were both heavy as in *\*'døø,man* above and *\*'her.ði.jooz* 'shepherds, nom pl' > *\*'her.ðooz*. Stresses also clashed in many other words where heavy syllables abutted even prior to syncope, e.g. *\*'wur.ðoo* 'words, nom/acc pl'. Riad proposes that stress clash was resolved by destressing the second syllable through shortening and other changes that left the affected syllable light: *\*'døø,man* > *'døø.ma*, *\*'her.ði.jooz* > Old Icelandic *hér.ðar*, *\*'wur.ðoo* > *'wor.ðu*, and *\*vål.jan* 'to choose' > *vel.jã*.

Riad hypothesizes that high F0 was one of the phonetic correlates of stress at this time, in addition to greater duration and intensity, and that this F0 peak survived destressing. Words that had undergone destressing would then have had two F0 peaks, one on the remaining primary stressed syllable and the other on the newly destressed syllable. After destressing, an F0 peak would no longer be a correlate of stress because only the first of these peaks would still be associated with a stressed syllable. Once destressing eliminated the prosodic equivalence of these two peaks, the first, the one still associated with the primary stressed syllable, was reanalyzed as a lexical H\* tone. This is the path of development to present-day



non-final stress bear Accent 1 in present-day Swedish, as in (1a) above. These words are the product of a variety of sound changes that added syllables to words that were once monosyllabic, notably the appearance of *svarabhakti* vowels before syllabic sonorants, e.g. *akr* > *aker* ‘field’, *fogl* > *fogel* ‘bird’, and *sookn* > *sokken* ‘parish’, all with Accent 1, and enclisis of the definite article, e.g. *and hinn* ‘duck the’ *and-in* > *anden*, also Accent 1 (1a), cf. *anden* ‘the ghost’ with Accent 2 (1b).

I will not take up here the debate as whether the contrast between word accents is privative or equipollent, nor if it’s privative whether Accent 2 is marked (Riad, 1998, 2003) or Accent 1 (Lahiri, Wetterlin, & Jönsson-Steiner, 2005), although as we will see it impinges on the development of *stød* in Danish, which is discussed in the next section.

## 5.2 Danish *stød*

Danish *stød* is phonetically a glottal stop or creaky voice that causes F0 to drop sharply at the end of the sonorant interval in a syllable bearing it. In examples here, it will be represented with an apostrophe following the end of this interval. A syllable can bear *stød* if it’s stressed, bimoraic (heavy), and its second mora is a vowel or sonorant. If the second mora is a sonorant consonant and word-final, then *stød* cannot occur. These conditions are referred to collectively as the “*stød* basis”. Syllables bearing *stød* are typically either the sole syllable in monosyllabic words or final stressed syllables in longer words. Following Itô & Mester (1997), Riad (2000) analyzes *stød* as the allophone of a general HL word contour that occurs when these conditions are met, i.e. when both tones are compressed into a single word-final stressed syllable. When the stressed syllable is non-final, the two tones can be distributed across two syllables and no *stød* occurs. Otherwise in the Scandinavian languages, only one tone can be aligned with the stressed syllable, Accent 2 can only occur when at least one unstressed syllable follows the stressed syllable, and only Accent 1 occurs on monosyllables and longer words with final stress. The only exceptions to these generalizations are a few northern dialects of Swedish where apocope of final unstressed vowels has led to Accent 2 being realized on a single syllable, in what is called “circumflex” accent. The syllable bearing

this accent is, however, lengthened to accommodate the two tones, so even these dialects attempt to stay within the one tone per syllable limit. The glottal stop/creaky voice and sharp drop in F0 that realize the L in *stød* can be interpreted as a similar accommodation to cramming the two tones into a single syllable.

Given the constraints on its distribution, it should not be surprising that Danish *stød* corresponds to Accent 1 in Swedish and Norwegian, which is similarly restricted in its distribution. The correspondence is not perfect, however, because *stød*'s distribution is constrained further by the *stød* basis, while Accent 1's distribution is only constrained by the position of stress in the word. If *stød* is simply the allophone of a general HL contour that is realized under the segmental and prosodic conditions specified by the *stød* basis, then, as Riad (2000) emphasizes, these conditions only determine how that contour can be realized phonetically and do not restrict whether a phonological object, here the HL contour, can occur, unlike the conditions on the distribution of Accent 1 in Swedish and Norwegian.

Riad (2000) proposes that the point of departure for the historical development of *stød* in Danish was a dialect quite like Central Stockholm Swedish; that is, a *stød* arose from word accents.<sup>26</sup> Specifically, *stød* arose from a source very much like connective Accent 2 in this Swedish dialect (2). The important characteristic of this accent for the development of *stød* is that the boundary L% is close to the stressed syllable of the compound's last constituent, and in this respect its F0 contour resembles Accent 1's. In non-compound forms bearing Accent 2, the L% tone is farther from the stressed syllable because the lexical H\* tone occurs there and the LH focal tone contour intervenes.

A present-day dialect spoken to the southwest of Stockholm in an area Riad refers to as Western Mälardalen, centered around the city of Eskilstuna, exemplifies two important intermediate steps in the development of *stød*. First, the tones in this dialects are shifted earlier in the word compared to Central Stockholm Swedish, such that the L% boundary

---

<sup>26</sup> This is precisely the opposite direction of historical development proposed by Liberman (1982) who treats a property very like *stød* as the source for the Swedish and Norwegian word accents. Lack of space prevents me from doing more here than mention this alternative.

tone is aligned with the stressed syllable, which crowds the tones of the focal contour earlier, too, displacing the focal L off the stressed syllable and aligning the focal H with it (3).

Stockholm $\begin{array}{c} \text{LHL} \\       \\ \text{s t o o r} \end{array}$	Eskilstuna $\begin{array}{c} \text{LHL} \\   \cdot \cdot \cdot \\ \text{s t o o r} \end{array}$	(3)
---	--	-----

(a) ‘big’ Accent 1

$\begin{array}{c} \text{H} \quad \quad \quad \text{L H} \quad \quad \text{L} \\   \quad \quad \quad   \quad \quad \quad   \\ \text{k o n ' c e e r n - , t r ä f f a r n a} \end{array}$	$\begin{array}{c} \text{H} \quad \quad \quad \text{H L} \\   \quad \quad \quad   \cdot \cdot \cdot \\ \text{k o n ' c e e r n - , t r ä f f a r n a} \end{array}$
--	---

(b) ‘the corporate meetings’ Accent 2

A consequence of this crowding of the focal H and L% boundary tones onto the stressed syllable is that in the appropriate pragmatic conditions, the L% boundary tone is realized as a steep drop in F0 that may end in creaky voice, a pronunciation described as *knorr* ‘curl’ or *knarr* ‘creak’ (!). The Eskilstuna curl is phonetically very similar to the Danish *stød*. The essential differences between them are that *stød* occurs in a narrower range of contexts than curl (those that meet the criteria of the *stød* basis), *stød* is categorical in contexts where it occurs rather than pragmatically conditioned, and word accents remain contrastive in Eskilstuna and Western Mälardalen generally, while they no longer contrast in Danish. The alignment of the L% boundary tone with the stressed syllable in the Eskilstuna-like ancestor of Danish motivated the segmental restrictions on the occurrence of *stød* that constitute the *stød* basis: the L tone needed sonorant material on which to realize itself.

The shift of the focal H tone to the stressed syllable in this Eskilstuna-like ancestor would have had the effect of aligning H tones with stressed syllables generally, which could lead to their reanalysis as correlates of stress and the loss of any phonological distinction between the lexical H\* of Accent 2 and the focal H. In words with Accent 1 or those bearing connective

Accent 2, nothing changes as a result of this reanalysis, but in words with Accent 2 that are neither compounds nor derived, the reanalysis of the focal H as a correlate of stress would have led speakers to hypothesize secondary stresses on the syllable bearing the focal H tone that were actually unstressed. At that point, none of the tones are lexical any longer.

Evidence supporting this last step can be found in the Danish phenomenon known as *slumrende* ‘slumbering’ *stød*, where a syllable that does not bear *stød* appears with it when a suffix is attached to the word, e.g. *sofa* [‘so:fa] ‘sofa’ and *sofaen* [‘so:fæ:’ən] ‘sofas’. There is no reason to expect the second syllable in this stem to bear a secondary stress, yet it must do so if it is to bear *stød*. What is striking about this account is that these sound changes effectively reverse the sequence of prosodic changes that created the word accents out of the F0 peaks aligned with stressed syllables in late Proto-Nordic.

This analysis appears to predict, incorrectly, that *stød* should correspond to Accent 2 in addition to Accent 1. Riad (2000) points to three properties that would have blocked such correspondences. First, the restrictions imposed by the *stød* basis would block many words with Accent 2 from developing *stød*. Second, in the absence of *slumrende stød* alternations, there would be no basis for positing *stød* in words ending in short vowels with Accent 2. Third, large numbers of vowel-final Danish words that correspond to Swedish or Norwegian words with Accent 2 end in schwa rather than a full vowel (unlike *sofa*), and schwa cannot bear *stød*. These characteristics all limit the number of possible Accent 2 words that could have met the conditions to develop *stød*.

### 5.3 Interim summary

In Riad’s (1998; 2003) account, word accents developed in the ancestor of Swedish and Norwegian once clash-resolving destressing of the second of two stressed syllables prosodically differentiated the H tone that remained aligned with the primary stressed syllable from that which had been aligned with the destressed syllable. The H that remained aligned was reanalyzed as a lexical H\* tone, producing Accent 2. Final primary stressed syllables would not have destressed, the H tone aligned them would not have been reanalyzed as a lexical

tone, and as a result Accent 1 words are not specified for tone lexically. Danish *stød* developed from word accents as a result of a shift of the L% boundary tone and the focal H tone to the stressed syllable, which led to a reanalysis of H tones as correlates of stress and the reintroduction of covert secondary stresses that reversed the sequence of sound changes that originally produced the word accents.

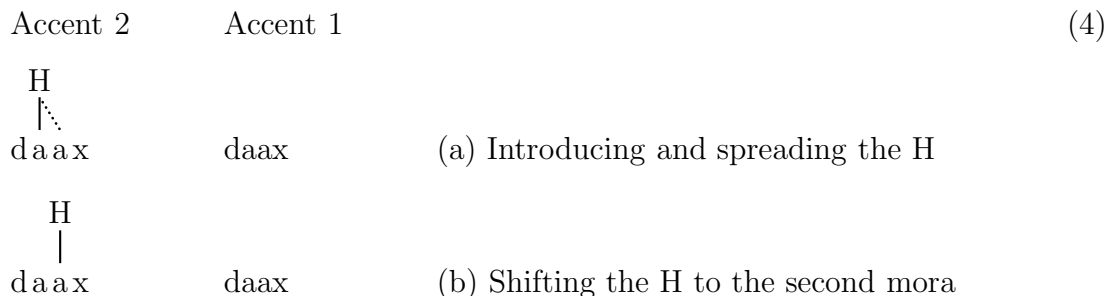
#### 5.4 Tonogenesis in Central Franconian

In Central Franconian, tones have developed as means of preserving the morphological contrasts between plural and singular forms of nouns, infinitive and bare stem forms of verbs, and dative singular and nominative and accusative singular forms of nouns and adjectives (Gussenhoven, 2000). Their development was made necessary by the combined effects of three prior sound changes (Table 14). First, the stem vowel in the plural, etc. underwent open syllable lengthening (OSL); second, the final schwa that carried the morphological distinction between these forms and the singular, etc. apocopated, and third, the stem vowels in the singular, etc. lengthened analogically to match the quantity of the stem vowels in the plural, etc. Up to this last change, the morphological contrasts had remained intact, but this analogical lengthening would have neutralized them if something were not done to compensate. What was done was first to extend the duration of an intonational H tone from the first to the second mora of the newly lengthened vowel in the singular, etc. (4a) and then to shift that H exclusively to the second mora (4b). The tone that developed on the singular, etc. is referred to as Accent 2, while the plural, etc. acquire Accent 1. Accent 2, also called the *Schleifton*, is phonetically long and high or rising, while Accent 1 or the *Stosston* is short, falling, and may end in creaky voice.

Stages	Singular	Plural
Middle High German (MHG)	dax	dayə
Open Syllable Lengthening	dax	da:ɤ
Apocope	dax	da:x
Analogical Lengthening	da:x	da:x

Table 14: Sequence of sound changes leading up to tonogenesis in Central Franconian.

The last step is the introduction, spreading, and shifting of the H tone to the second mora in the singular, etc. (4).



For lack of space, I will not discuss the subsequent changes in the F0 value of tones in particular Franconian dialects nor the changes in the timing of the introduced tone relative to the timing of the tones introduced by the intonational component (see Gussenhoven, 2000, for a particularly insightful treatment). Instead, I will take up the correlations between this development and two segmental properties of the syllables bearing these tones, because they obviously connect to the discussion elsewhere in this paper of the segmental sources for tones. Accent 1 tends to occur on syllables with non-high vowels [e:, ø:, o:, æ:, a:] in MHG, while Accent 2 tends to occur on monosyllables with high vowels [i:, y:, u:] and diphthongs [ɛi, œy, ʌu] in MHG, and Accent 2 tends to occur on syllables followed by a voiceless obstruent in the onset of the next syllable. Both correlations have led to proposals that one or the other of these segmental properties was the historical source of the tones, which Gussenhoven (2000) shows cannot be so.

Non-high vowels differ in two potentially relevant ways from high vowels: they are longer and have lower intrinsic F0. The length correlation is, however, backward, as the longer non-high vowels cooccur with the shorter Accent 1. The F0 correlation is in the right direction, high vowels and diphthongs with high offglides cooccur with Accent 2 which ends in a H tone. However, there is no evidence that vowels have changed height or that height contrasts have merged as a result of tonogenesis in Central Franconian, so the most one can say is that accents cooccur with vowels that would have similar effects on F0. The cooccurrence

of Accent 2 with syllables followed by a voiceless obstruent is similar. As described in detail above, F0 is typically higher at the beginning of vowels following voiceless obstruents than those following voiced ones. This raising is compatible with the high or rising F0 of Accent 2, but it occurs on the following not the preceding vowel, and voiceless obstruents have not otherwise been shown to produce high tones on preceding vowels. Therefore, the association between Accent 2 and following voiceless obstruents, too, remains no more than a correlation between similar raising effects on F0 and cannot be the historical source for Accent 2 (see Gussenhoven, 2000, for further arguments along these lines). Finally, neither of these correlations amounts to a condition on the occurrence of either of the accents, so they are unlike the *stød* basis.

## 5.5 Summary

Word accents developed in the Scandinavian languages as a result of clash-resolving de-stressing and the resulting reinterpretation of the F0 peaks that had previously been predictable correlates of stress. *Stød* developed subsequently in Danish as a result of a shift of tones to earlier positions that realigned F0 peaks with stressed syllables and reversed the reanalysis that had led to the original development of the word accents. In Central Franconian, the motivation for tonogenesis was quite different, preservation of a number of morphological contrasts that would have been neutralized by analogical lengthening. The original development and present-day distributions of the word accents in the Scandinavian languages are certainly morphologically conditioned, but this conditioning is a residue of the stresses borne by inflectional and derivational suffixes and by the constituents of compounds, and not an attempt to preserve morphological contrasts. There is little reason therefore to treat tonogenesis in Central Franconian as related in any way to that in Scandinavian.

## 6 Summary and conclusions

In this chapter, I have reviewed the ways and means that languages have acquired distinct tone. Considerable attention was paid to tonogenesis and tone splitting in a number of

language families in East and Southeast Asia, including Sino-Tibetan, Hmong, Tai, Kam-Sui, and to those few Mon-Khmer and Chamic languages that have also developed tone. Tones have developed and multiplied in these languages, as well as in Athabaskan because the laryngeal articulations of consonants alter or perturb F0 in neighboring vowels. Two other languages in which tone has developed from this source were also discussed, Jabem and Korean. Just two languages, U and Lugbara, have apparently split their tones under the influence of intrinsic F0 differences between vowels contrasting for height and ATR, respectively. In all these cases, a contrast may ultimately be transferred from the segments that originally carried it, consonants contrasting in their laryngeal articulations or vowels contrasting for height or ATR, to F0 differences on the vowels. These are thus all cases in which the formerly predictable effects on F0 of adjacent consonants or the vowel itself have been phonologized. The chapter closed with discussion of how the word accents developed in Swedish and Norwegian, as a result of clash-resolving destressing and the resulting reanalysis of F0 peaks as lexical and focal tones rather than as correlates of stress, how *stød* subsequently developed from these word accents in Danish when F0 peaks realigned with stressed syllables and were once again analyzable as stress correlates, and how tone developed to preserve morphological contrasts threatened by analogical lengthening in Central Franconian. The Central Franconian case can also be described as a transfer of contrast, but in this case, the phonetic material from which the tones came is entirely independent of the prior changes that prompted the introduction of tones, although high final F0 of Accent 2 in Central Franconian does correlate with the raising of F0 by the high vowels and voiceless obstruents in the words bearing this accent. The word accents in Swedish and Norwegian are a phonologization of what was once a predictable correlate of stress. They could be interpreted as a transfer of contrast in that Accent 2 emerged in words in which a secondary stresses were lost to resolve stress clash, and Accent 1 emerged on words in which no stress clash occurred because the final syllable bore primary stress. Subsequent sound changes affecting Accent 1 words introduced following unstressed syllables, and words with non-final

primary stress now contrast for word accent.

In both Athabaskan and East and Southeast Asian languages, it is necessary to recognize that consonants with a particular laryngeal articulation can be the source of higher tones in one language and lower tones in another. In both cases, this variability can probably be attributed to the presence in these languages of contrasts for laryngeal articulations in sonorants as well as obstruents. Unlike the laryngeal articulations of obstruents, there are no phonetic constraints on laryngeal articulations in sonorants that prevent them from being pronounced so as to either raise or lower F0. Individual languages, groups of languages, or whole families opt to pronounce the laryngeal articulations in sonorants one way or the other, and the obstruents with the same laryngeal articulations follow suit. The essential difference between the Athabaskan versus East and Southeast Asian cases is that in Athabaskan the laryngeal contrasts have been preserved in the sonorants and lost in the obstruents, while in East and Southeast Asia, contrasts are transferred more thoroughly from the laryngeal articulations of sonorants than obstruents, which may retain some of their laryngeal articulations. The result of this difference is that particular groups of languages and individual languages in the Athabaskan family have repeatedly changed their pronunciation of the sonorants, while in East and Southeast Asia, there is far more uniformity within language groups, with the exception of the Voiced Low versus Voiced High difference. These facts suggest that consonants should be specified for tone, using the features [stiff] and [slack], in addition to [voice], [spread glottis], and [constricted glottis].

There are at least three quite obvious gaps in this chapter's coverage, the origins of tone in the many tonal Central American, African, and New Guinean languages (except Jabem). Tones have been reconstructed for the proto-languages from which the present-day Central American and African languages descend, so it has not been possible to work out how these languages became tonal.<sup>27</sup> Catalogues are beginning to emerge of the types of tone languages

---

<sup>27</sup> This is not to say of course that tones have not changed between the proto-languages and the present-day, but the variety of those changes is so bewildering that I could not present them succinctly here.

found in New Guinea, particularly in the very large Trans-New Guinea stock (Donohue, 1997, 2005; Ross, 2005), but again the data present such a blizzard of varieties that it's beyond what I could attempt here to draw generalizations about their historical development.

## References

- Alderete, J. (2001). *Morphologically governed accent in optimality theory*. New York: Routledge.
- Andersen, T. (1986). Tone splitting and vowel quality: Evidence from Lugbara. *Studies in African Linguistics*, 17, 55-68.
- Avery, P., & Idsardi, W. J. (2001). Laryngeal dimensions, completion, and enhancement. In T. A. Hall (Ed.), *Distinctive feature theory* (p. 41-70). Berlin: Mouton de Gruyter.
- Bell-Berti, F., & Hirose, H. (1975). Palatal activity in voicing distinctions: A simultaneous fiberoptic and electromyographic study. *Journal of Phonetics*, 3, 69-74.
- Blevins, J. (2004). *Evolutionary phonology: The emergence of sound patterns*. Cambridge: Cambridge University Press.
- Bradshaw, J. (1979). Obstruent harmony and tonogenesis in Jabêm. *Lingua*, 49, 189-205.
- Brown, J. M. (1975). The great tone split: Did it work in two different ways? In J. G. Harris & J. R. Chamberlin (Eds.), *Studies in Tai linguistics in honor of William J. Gedney* (p. 33-48). Bangkok: Office of State Universities.
- Bruce, G. (1977). *Swedish word accents in sentence perspective*. Lund: Gleerup.
- Bruce, G. (1999). Word tone in Scandinavian languages. In H. van der Hulst (Ed.), *Word prosodic systems in the languages of Europe* (p. 606-633). Berlin: Mouton de Gruyter.
- Bruce, G., & Gårding, E. (1978). A prosodic typology for Swedish dialects. In E. Grding, G. Bruce, & R. Bannert (Eds.), *Nordic prosody, travaux de l'institut de linguistique de lund xiii* (p. 219-228). Lund: Gleerups.
- Cassimjee, F. (1998). *Isixhosa tonology: An optimal domains theory analysis*. München: LINCOM Europa.

- Cassimjee, F., & Kisseberth, C. W. (1998). Depressor consonants in Mijikenda and Nguni. *Proceedings of Annual Meeting of the Berkeley Linguistics Society*, 18, 26-40.
- Collier, R. (1975). Physiological correlates of intonation patterns. *Journal of the Acoustical Society of America*, 58, 249-55.
- Connell, B. (2002). Tone languages and the universality of intrinsic f<sub>0</sub>: Evidence from Africa. *Journal of Phonetics*, 30, 101-129.
- Denning, K. (1989). *The diachronic development of phonological voice quality with special reference to Dinka and other Nilotic languages*. Unpublished doctoral dissertation, Stanford University.
- Diffloth, G. (1989). Proto-Austroasiatic creaky voice. *Mon-Khmer Studies*, 15, 139-154.
- Donnelly, S. (2009). Tone and depression in Phuthi. *Language Sciences*, 31, 161-178.
- Donohue, M. (1997). Tone in New Guinea. *Linguistic Typology*, 1, 347-386.
- Donohue, M. (2005). Tone and the Trans New Guinea languages. In S. Kaji (Ed.), *Cross-linguistic studies of tonal phenomena: Historical development, tone-syntax interface, and descriptive studies* (p. 33-54). Tokyo: Research Institute for Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.
- Downer, G. B. (1959). Derivation by tone-change in Classical Chinese. *Bulletin of the School of African and Oriental Studies*, 22, 258-290.
- Downing, L., & Gick, B. (2005). Voiceless tone depressors in Nambya and Botswana Kalang'a. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society*, 27, 65-80.
- Downing, L., & Schadeberg, T. (2009). On pitch lowering not linked to voicing: Nguni and Shona group depressors. *Language Sciences*, 31, 179-198.

- Edmondson, J. A., & Esling, J. H. (2006). The valves of the throat and their functioning in tone, vocal register and stress: Laryngographic case studies. *Phonology*, *23*, 157-191.
- Erickson, D., Honda, K., Hirai, H., Beckman, M. E., & Seiji, N. (1994). Global pitch range and the production of low tones in english intonation. In *Third International Conference on Spoken Language Processing-1994* (p. 651-654). Yokohama.
- Ewan, W. G. (1976). *Laryngeal behavior in speech*. Unpublished doctoral dissertation, University of California, Berkeley.
- Gårding, E. (1977). *The Scandinavian word accents*. Lund: Gleerup.
- Gregerson, K. J. (1976). Tongue-root and register in Mon-Khmer. In P. Jenner, L. Thompson, & S. Starosta (Eds.), *Austro-asiatic studies* (Vol. 1, p. 323-369). Honolulu: University of Hawaii Press.
- Gussenhoven, C. (2000). On the origin and development of the Central Franconian tone contrast. In A. Lahiri (Ed.), *Analogy, leveling, and markedness* (p. 215-260). Berlin: Mouton de Gruyter.
- Gussenhoven, C., & Bruce, G. (1999). Word prosody and intonation. In H. van der Hulst (Ed.), *Word prosodic systems in the languages of Europe* (p. 233-271). Berlin: Mouton de Gruyter.
- Halle, M., & Stevens, K. N. (1971). A note on laryngeal features. *Research Laboratory in Electronics, Quarterly Progress Report*, *101*, 198-213.
- Haudricourt, A.-G. (1954a). Comment reconstruire le chinois archaïque. *Word*, *10*, 351-364.
- Haudricourt, A.-G. (1954b). De l'origine ds tons en vietnamien. *Journal Asiatique*, *242*, 69-82.

- Haudricourt, A.-G. (1972). Two-way and three-way splitting of tonal systems in some Far Eastern languages. In J. G. Harris & B. Noss Richard (Eds.), *Tai phonetics and phonology* (p. 58-86). Bangkok: Central Institute of English Language.
- Henderson, E. J. A. (1967). Grammar and tone in South-East Asian languages. *Wissenschaftliche Zeitschrift der Karl-Marx-Universität-Leipzig (Gesellschafts- und Sprachwissenschaftliche Reihe), Part 1/2*, 171-178.
- Hermans, B. (1999). A description of tonal accent in a Limburgian dialect. In H. van der Hulst (Ed.), *Word prosodic systems in the languages of Europe* (p. 633-658). Berlin: Mouton de Gruyter.
- Hirose, H., Lee, C. Y., & Ushijima, T. (1974). Laryngeal control in Korean stop production. *Journal of Phonetics*, 2, 145-152.
- Hirose, H., Park, H. S., Yoshioka, H., Sawashima, M., & Umeda, H. (1981). An electromyographic study of laryngeal adjustments for the Korean stops. *Annual Bulletin, Research Institute of Logopedics and Phoniatrics, University of Tokyo*, 15, 31-43.
- Hombert, J.-M. (1977). Consonant types, vowel height, and tone in Yoruba. *Studies in African Linguistics*, 8, 173-190.
- Hombert, J.-M. (1978). Consonant types, vowel quality, and tone. In V. Fromkin (Ed.), *Tone: A linguistic survey* (p. 77-111). New York: Academic Press.
- Hombert, J.-M., Ohala, J. J., & Ewan, W. G. (1979). Phonetic explanations for the development of tones. *Language*, 55, 37-58.
- Honda, K., Hirai, H., Masaki, S., & Shimada, Y. (1999). Role of vertical larynx movement and cervical lordosis in f0 control. *Language and Speech*, 42, 401-411.
- Huffman, F. (1976). The register problem in fifteen Mon-Khmer languages. In P. N. Jenner,

- L. C. Thompson, & S. Starosta (Eds.), *Austroasiatic studies* (Vol. 1, p. 575-589). Honolulu: University of Hawai'i Press.
- Hyman, L. M. (2008). Enlarging the scope of phonologization. *UC Berkeley Phonology Lab Annual Report*, 382-409.
- Itô, J., & Mester, A. (1997). *Stø?det i dansk*. Handout, Scandinavian Summer School in Generative Phonology, Hvalfjarðarströnd.
- Jessen, M., & Roux, J. (2002). Voice quality differences associated with stops and clicks in Xhosa. *Journal of Phonetics*, 30, 1-52.
- Jun, S.-A. (1996). Influence of microprosody on macroprosody: A case of phrase-initial strengthening. *University of California, Los Angeles, Working Papers in Phonetics*, 92, 97-116.
- Kagaya, R. (1974). A fiberscopic and acoustic study of the Korean stops, affricates, and fricatives. *Journal of Phonetics*, 2, 161-180.
- Kingston, J. (1985). *Phonetics and phonology of the timing of oral and glottal events*. Ph.d. dissertation, University of California, Berkeley.
- Kingston, J. (1990). Articulatory binding. In J. Kingston & M. Beckman (Eds.), *Papers in laboratory phonology i* (p. 406-434). Cambridge: Cambridge University Press.
- Kingston, J. (1991). Integrating articulations in the perception of vowel height. *Phonetica*, 48, 149-179.
- Kingston, J. (1992). The phonetics and phonology of perceptually motivated articulatory coordination. *Language and Speech*, 35, 99-113.
- Kingston, J. (2005). The phonetics of Athabaskan tonogenesis. In S. Hargus & K. Rice (Eds.), *Athabaskan prosody* (p. 137-184). Amsterdam: John Benjamins.

Kingston, J., & Diehl, R. L. (1994). Phonetic knowledge. *Language*, 70, 419-454.

Kingston, J., & Diehl, R. L. (1995). Intermediate properties in the perception of distinctive feature values. In B. Connell & A. Arvaniti (Eds.), *Phonology and Phonetics: Papers in Laboratory Phonology* (Vol. IV, p. 7-27). Cambridge: Cambridge University Press.

Kingston, J., Diehl, R. L., Kirk, C. J., & Castleman, W. A. (2008). On the internal perceptual structure of distinctive features: The [voice] contrast. *Journal of Phonetics*, 36(1), 28-54.

Kingston, J., & Solnit, D. (1989). The inadequacy of underspecification. *Proceedings of North East Linguistic Society*, 19, 264-278.

Krauss, M. E. (2005). Athabaskan tone. In S. Hargus & K. Rice (Eds.), *Athabaskan prosody* (p. 51-136). Amsterdam: John Benjamins.

Ladd, D. R., & Silverman, K. E. A. (1984). Vowel intrinsic pitch in connected speech. *Phonetica*, 41, 31-40.

Lahiri, A., Riad, T., & Jacobs, H. (1999). Diachronic prosody. In H. v. d. Hulst (Ed.), *Word prosodic systems in the languages of europe* (p. 335-421). New York: Mouton de Gruyter.

Lahiri, A., Wetterlin, A., & Jönsson-Steiner, E. (2005). Lexical specification of tone in North Germanic. *Nordic Journal of Linguistics*, 28, 61-96.

Laver, J. (1980). *The phonetic description of voice quality*. Cambridge, UK: Cambridge University Press.

Leer, J. (1979). *Proto-Athabaskan verb stem variation, part one: Phonology* (Vol. 1). Fairbanks: University of Alaska, Alaska Native Language Center.

Leer, J. (1999). Tonogenesis in Athabaskan. In S. Kaji (Ed.), *Cross-linguistic studies of tonal phenomena, tonogenesis, typology, and related topics* (p. 37-66). Tokyo: Institute of

the Study of the Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.

Leer, J. (2001). Shift of tonal markedness in Northern Tlingit and Southern Athabaskan. In S. Kaji (Ed.), *Cross-linguistic studies of tonal phenomena: Tonogenesis, Japanese accentology, and other topics* (p. 61-89). Tokyo: Institute of the Study of the Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.

Li, F.-K. (1948). The distribution of initials and tones in Sui language. *Language*, 24, 160-167.

Li, F.-K. (1965). The Tai and Kam-Sui languages. *Lingua*, 14, 148-179.

Li, F.-K. (1977). *A handbook of Comparative Tai* (Vol. 15). Honolulu: University of Hawai'i Press.

Liberman, A. (1982). *Germanic accentology, v. 1: The Scandinavian languages*. Minneapolis: University of Minnesota Press.

Lindau, M. (1979). The feature expanded. *Journal of Phonetics*, 7, 163-176.

Löfqvist, A., Baer, T., McGarr, N., & Seider Story, R. (1989). The cricothyroid muscle in voicing control. *Journal of the Acoustical Society of America*, 85, 1314-1321.

Löfqvist, A., & Yoshioka, H. (1980). Laryngeal activity in Swedish obstruent clusters. *Journal of the Acoustical Society of America*, 68, 792-801.

Maddieson, I. (1978). Universals of tone. In J. Greenberg (Ed.), *Universals of human language, vol. 2: Phonology* (p. 335-365). Stanford: Stanford University Press.

Maddieson, I. (1984). The effects on f<sub>0</sub> of a voicing distinction in sonorants and their implications for a theory of tonogenesis. *Journal of Phonetics*, 12, 9-15.

- Maddieson, I. (2003). The sounds of Bantu languages. In D. Nurse & G. Philippson (Eds.), *The Bantu languages* (p. 15-41). London: Routledge.
- Maddieson, I., & Pang, K.-F. (1993). Tone in Utsat. *Oceanic Linguistics Special Publications*, 24, 75-89.
- Matisoff, J. A. (1972). *The Loloish tonal split revisited* (Vol. 7). Berkeley: University of California.
- Matisoff, J. A. (1973). Tonogenesis in Southeast Asia. In L. M. Hyman (Ed.), *Consonant types and tone* (p. 97-117). Los Angeles: University of Southern California.
- Matisoff, J. A. (1979). Problems and progress in Lolo-Burmese: Quo vadimus? *Linguistics of the Tibeto-Burman area*, 4, 11-43.
- Mazaudon, M. (1977). Tibeto-Burman tonogenetics. *Linguistics of the Tibeto-Burman Area*, 3.2, 1-123.
- Mazaudon, M. (1978). Consonantal mutation and tonal split in the Tamang sub-family of Tibeto-Burman. *Kailash*, 6.3, 157-179.
- Mazaudon, M. (2009). Paths to tone in the Tamang branch of Tibeto-Burman (Nepal). In G. de Vogelaer & G. Seiler (Eds.), *Dialects as a testing ground for theories of language change*. Amsterdam: John Benjamins.
- Mazaudon, M., & Michaud, A. (2008). Tonal contrasts and initial consonants: A case study of Tamang, a 'missing link' in tonogenesis. *Phonetica*, 65, 231-256.
- Mei, T.-l. (1970). Tones and prosody in Middle Chinese and the origin of the Rising tone. *Harvard Journal of Asian Studies*, 30, 86-100.
- Mei, T.-l. (1980). Si sheng bie yi zhong de shijian cengci. *Zhongguo Yuwen*, 6, 427-443.

- Moreton, E. (2004). Non-computable functions in optimality theory. In J. J. McCarthy (Ed.), *Optimality theory in phonology* (p. 141-164). Oxford: Blackwell Publishing.
- Moreton, E. (2008a). Learning bias as a factor in phonological typology. In C. Chang & A. Haynie (Eds.), *Proceedings of the 26th Meeting of the West Coast Conference on Formal Linguistics* (p. 393-401). Somerville, MA: Cascadilla Press.
- Moreton, E. (2008b). Modeling modularity bias in phonological pattern learning. In N. Abner, J. Bishop, & K. Ryan (Eds.), *Proceedings of the 27th Meeting of the West Coast Conference on Formal Linguistics* (p. 1-16). Somerville, MA: Cascadilla Press.
- Moreton, E. (to appear). Underphonologization and modularity bias. In S. Parker (Ed.), *Phonological argumentation: Essays on evidence and motivation*. London: Equinox.
- Ohala, J. J. (1973). The physiology of tone. In L. M. Hyman (Ed.), *Consonant types and tone* (Vol. 1, p. 1-14).
- Ohala, J. J., & Eukel, B. (1987). Explaining the intrinsic pitch of vowels. In R. Channon & L. Shockey (Eds.), *In honour of Ilse Lehiste* (p. 207-215). Dordrecht: Foris.
- Ohala, J. J., & Riordan, C. J. (1979). Passive vocal tract enlargement during voiced stops. In J. J. Wolf & D. H. Klatt (Eds.), *Speech communication papers* (p. 89-92). New York: Acoustical Society of America.
- Phu, V. H., Edmondson, J. A., & Gregerson, K. J. (1992). East Cham as a tone language. *Mon-Khmer Studies*, 20, 31-43.
- Pulleyblank, E. G. (1962). The consonantal system of Old Chinese. *Asia Major, Series 2*, 9, 58-144 and 206-265.
- Reinholt Peterson, N. (1978). Intrinsic fundamental frequency of danish vowels. *Journal of Phonetics*, 6, 177-189.

- Riad, T. (1998). The origin of Scandinavian tone accents. *Diachronica*, 15, 63-98.
- Riad, T. (2000). The origin of the Danish *stød*. In A. Lahiri (Ed.), *Analogy, leveling, and markedness* (p. 261-300). Berlin: Mouton de Gruyter.
- Riad, T. (2003). Diachrony of the Scandinavian accent typology. In P. Fikkert & H. Jacobs (Eds.), *Development in prosodic systems* (p. 91-144). Berlin: Mouton de Gruyter.
- Ross, M. D. (1993). Tonogenesis in the North Huon Gulf Chain. *Oceanic Linguistics Special Publications*, 24, 133-153.
- Ross, M. D. (2005). Towards a reconstruction of the history of tone in the Trans New Guinea family. In S. Kaji (Ed.), *Cross-linguistic studies of tonal phenomena: Historical development, tone-syntax interface, and descriptive studies* (p. 3-31). Tokyo: Research Institute for Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.
- Rycroft, D. K. (1980). *The depression feature in Nguni languages and its interaction with tone* (Vol. 8). Grahamstown: Department of African Languages, Rhodes University.
- Rycroft, D. K. (1983). Tone patterns in Zimbabwean Ndebele. *Bulletin of the School of Oriental and African Studies, University of London*, 46, 77-135.
- Sagart, L. (1999). The origin of Chinese tones. In S. Kaji (Ed.), *Cross-linguistics studies of tonal phenomena: Tonogenesis, typology, and related topics* (p. 91-103). Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.
- Schachter, P. (1976). An unnatural class of consonants in Siswati. *Studies in African Linguistics, Supplement 6*, 211-220.
- Schuessler, A. (1985). *The function of Qusheng in early Zhou Chinese* (Vol. 87). Canberra.

- Schuh, R. (1978). Tone rules. In V. Fromkin (Ed.), *Tone: A linguistic survey* (p. 221-256). New York: Academic Press.
- Silva, D. J. (2006). Acoustic evidence for the emergence of tonal contrast in contemporary Korean. *Phonology*, 23, 287-308.
- Solnit, D., & Kingston, J. (1988). Voiced high splits: Putting the blame on sonorants. In *63rd Annual Meeting of the Linguistic Society of America*. New Orleans.
- Steele, S. A. (1986). Interaction of vowel f0 and prosody. *Phonetica*, 43, 92-105.
- Stevens, K., & Blumstein, S. (1981). The search for invariant acoustic correlates of phonetic features. In J. Miller & P. Eimas (Eds.), *Perspectives on the study of speech* (p. 1-38). New Jersey: Earlbaum.
- Strazny, P. (2003). Depression in Zulu: Tonal effects of segmental features. In J. van de Weijer, V. J. van Heuven, & H. van der Hulst (Eds.), *The phonological spectrum: Segmental structure* (Vol. 1, p. 223-239). Amsterdam: John Benjamins.
- Suwilai, P. (2003). Khmu dialects: A case of register complex and tonogenesis. In S. Kaji (Ed.), *Cross-linguistic studies of tonal phenomena: Historical development, phonetics of tone, and descriptive studies* (p. 13-28). Tokyo: Institute of the Study of the Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.
- Svantesson, J.-O. (1989). Tonogenetic mechanisms in northern Mon-Khmer. *Phonetica*, 46, 60-79.
- Thongkum, T. L. (1987). Phonation types in Mon-Khmer languages. *University of California, Los Angeles, Working Papers in Phonetics*, 76, 29-48.
- Thurgood, G. (1999). *From Ancient Cham to modern dialects: Two thousand years of language contact and change* (Vol. 28). Honolulu: University of Hawai'i Press.

- Thurgood, G. (2002). Vietnamese and tonogenesis. *Diachronica*, 19, 333-363.
- Truill, A. (1990). Depression without depressors. *South African Journal of African Languages*, 10, 166-172.
- Truill, A., Khumalo, J. S. M., & Fridjhon, P. (1987). Depressing facts about Zulu. *African Studies*, 46, 255-274.
- Westbury, J. R. (1983). Enlargement of the supraglottal cavity and its relation to stop consonant voicing. *Journal of the Acoustical Society of America*, 73, 1322-1336.
- Whalen, D. H., Gick, B., Kumada, M., & Honda, K. (1999). Cricothyroid activity in high and low vowels: exploring the automaticity of intrinsic f<sub>0</sub>. *Journal of Phonetics*, 27, 125-142.
- Whalen, D. H., & Levitt, A. G. (1995). The universality of intrinsic f<sub>0</sub> of vowels. *Journal of Phonetics*, 23, 349-366.
- Yoshioka, H., Löfqvist, A., & Hirose, H. (1981). Laryngeal adjustments in the production of consonant clusters and geminates in American English. *Journal of the Acoustical Society of America*, 70, 1615-1623.
- Zee, E. (1980). Tone and vowel quality. *Journal of Phonetics*, 8, 247-258.