THE PHONETICS OF ATHABASKAN TONOGENESIS

JOHN KINGSTON
University of Massachusetts

1. *Introduction*

Syllables contrast for high vs low tone in many Athabaskan languages. The high tones in some of these languages correspond to low tones in others, and vice versa. The principal goal of this article[1] is to explain how the tones in some

---

[1] This article is a thorough reworking of an unpublished manuscript "The phonetics and phonology of Athabaskan tonogenesis" (1985), itself extracted from my dissertation *The phonetics and phonology of the timing of oral and glottal events* (also 1985). I am very grateful to Sharon Hargus and Keren Rice for giving me the opportunity to wrestle with these problems again. I also appreciate their guiding me to relevant work on Athabaskan that has come out since that earlier manuscript was written, and commenting thoroughly and insightfully on earlier versions of this article. Robert Hagiwara insightfully criticized §3, which is now much more readable because of his efforts. The original manuscript as well as the present one has been profoundly shaped by the work on Athabaskan tone of Michael Krauss and Jeff Leer, without which no version of this manuscript could have been written. Of great use, too, has been Matthew Gordon's article on the phonetics and phonology of glottal contrasts in stem-final consonants in Hupa, which he kindly dispatched to me at very short notice. I have also benefitted from seeing an advance copy of Gary Holton's paper in this volume. At short notice, too, Victor Golla and Michael Krauss filled in gaps in my knowledge about the retention of stem-final glottalic consonants in Pacific Coast Athabaskan languages. For help in evaluating the explanations for recent tone reversals, I thank the audience at the tonogenesis workshop organized by Aditi Lahiri at the Schloß Freudental, Konstanz, 20-23 March 2002. Finally, one of the two central ideas of the original manuscript, that tones but not other features could reverse their values, was developed during a very useful conversation with Victor Golla and Kenneth Whistler. In this article, I now reject that idea and replace it with what I think are more plausible alternatives, but that conversation keeps its value, as a foil against which to measure the

Athabaskan languages came to have the opposite values from those in others. The historical facts indicate that these mirror-image correspondences developed in two quite different ways, and at two quite different times in the history of the family.

Some Athabaskan languages with opposite tones have no more recent common ancestor than Proto-Athabaskan (PA), e.g. Navajo and Chipewyan. In this article, I propose that these distantly related languages descend from different dialects of PA, whose speakers pronounced differently the stem-final glottalic consonants from which tone originally evolved. Other Athabaskan languages with opposite tones have a much more recent common ancestor, i.e. low tone in Dogrib corresponds to high tone in Slave, Hare, Mountain, Bearlake, and Chipewyan within the Mackenzie River subgroup, high tone in Tanacross corresponds to low tone in Upper and Lower Tanana within the Tanana subgroup, and high tone in Northern Tutchone corresponds to low tone in Southern Tutchone. In each of these cases, the languages are closely related, and to a large extent mutually intelligible (Krauss, 1979, this volume; Leer, 1999). Because these closely related languages or dialects have split apart quite recently, a tone contrast had long since nearly completely replaced the Proto-Athabaskan contrast between stems ending in glottalic vs non-glottalic consonants in their common ancestor, and many of the opposite tones could no longer have developed directly from different pronunciations of these consonants. My proposal for these languages is that the tones that originally developed from the Proto-Athabaskan contrast between stem-final glottalic vs non-glottalic consonants recently reversed their values.

This article is organized as follows. In §2, I briefly describe the linguistic split between tonal and non-tonal Athabaskan languages and then the conditions for the evolution of tone from stem-final glottalic consonants. That presentation is aimed at this article's second goal: explaining phonetically how tonogenesis depends on the manner of articulation of the stem-final consonant and on the duration of the stem's nucleus. Then in §3 I return to the article's first goal, where I argue that both high and low tone could have evolved directly from different pronunciations of stem-final glottalic consonants. In §4, I show how the more recent reversals between high and low tone could have arisen in one of two ways. First, despite bearing contrastive tones, some syllables in many present-day tone languages also still differ in whether they end in a glottal stop or glottalic sonorant. The contrast between glottalic and non-glottalic final obstruents has long since been replaced by tone in

---

new proposals. Conversation with Joyce McDonough gave me one of the key pieces of the new proposal: that stem and prefix tones are likely to differ. I thank all these people for their generous help; none is responsible for any errors that may remain.

these languages. If the glottalic articulation were pronounced differently in different languages like all the glottalic final consonants had been earlier, then tone values could have recently reversed between even closely related languages in the same way that they had earlier. Alternatively, speakers' late realization of F0 targets in shortened syllables could have led listeners to mistake one syllable's tone for the preceding syllable's. If the preceding syllable's tone were opposite in value, this perceptual mistake would reverse the current syllable's tone value. These two ways of getting the recent tone reversals resemble one another in that both depend on analogy to extend their effects throughout the vocabulary. §5 contains brief concluding remarks.

2. ***The distribution of tone in Athabaskan***

2.1 *The history and geography of tone in Athabaskan*

The geographical distribution of tone in the Athabaskan languages divides the family in four: (i) the Pacific Coast languages, which don't distinguish syllables for tone, (ii) the non-tonal languages of southern and western Alaska and Western British Columbia, (iii) the tonal Northern Athabaskan languages of northern and eastern Canada and northern and eastern Alaska, and (iv) the tonal Apachean languages (Shipley, 1978; Krauss & Golla, 1981; Young, 1983; Goddard, 1996). However, tonal developments support two or at most three subgroups, which don't correspond to these geographical units: one in which tone didn't evolve from stem-final glottalic consonants and another in which it did, with the latter subdivided into groups whose speakers pronounced the stem-final glottalic consonants so as to raise or lower F0 on preceding vowels. Tone evolved via a shift of the glottalic articulation of the stem-final consonant to the preceding vowel in the form of a distinctive non-modal, 'constricted' voice quality. I propose that the contrast shifted from the stem-final consonant to the preceding vowel after the ancestors of both the Pacific Coast and non-tonal Alaskan languages separated linguistically from the ancestors of what would become the tonal Athabaskan languages. The simplest hypothesis is that this shift happened once, in a single protodialect that is ancestral to all the tonal languages.[2] For a comprehensive list of whch languages underwent

---

[2] No other sound changes are shared by the daughter languages that replaced the glottalic vs non-glottalic contrast in stem-final consonants with tones, nor by those that did not. Nonetheless, the complex conditioning of tonogenesis by all the properties of the rime discussed below is shared by all the tonal daughters, so it is probably more than just convenient to treat the tonal daughters as arising from just one protodialect, in which the

each of the developments described above, see Krauss (1979, this volume;see also Table 3 in Goddard, 1996, p. 5, for a complete list of Athabaskan languages).

If tone replaces a stem-final contrast between glottalic and non-glottalic consonants, then these two properties are expected to be mutually exclusive. Their distribution largely accords with this expectation: the Pacific Coast languages and most of the nontonal Alaskan languages retain the stem-final consonantal contrast and their vowels don't contrast for tone, while very nearly all of the tonal Canadian and Alaskan Athabaskan and the Apachean languages have lost the stem-final consonantal contrast and their vowels do contrast for tone. The only exceptions are those most western Alaskan languages, Deg Hit'an (Ingalik), Holikachuk, and some dialects of Koyukon,[3] and the southwest Canadian language, Babine, which have lost the stem-final consonantal contrast and also lack a tone contrast on their vowels, and Sarcee, the most southeastern Canadian Athabaskan language, in which glottalic and non-glottalic consonants contrast stem-finally and vowels also contrast independently for tone (Krauss, 1979, this volume; Cook, 1984).

All the exceptions may reflect the outlying positions of these languages. Deg Hit'an, Holikachuk, and toneless dialects of Koyukon could simply have lost the stem-final contrast after the western Alaskan languages had separated from the tonal protodialect, and thus did not undergo the shift of the contrast that took place in that protodialect. Because it is closely related to neighboring tonal languages, Babine, at the southwestern edge of Canadian Athabaskan, probably did develop tone originally from the stem-final contrast but has since lost it. Sarcee, at the southeastern edge, did not entirely lose the stem-final glottalic:non-glottalic contrast once tone had evolved; it survives in the perfective allomorph of reduced vowel verbs and in the perfective allomorph of full vowel verbs when they are followed by a vowel-initial suffix (Cook, 1984). (The reduced:full contrast between vowels is discussed immediately below.)

In this article, I explain how the languages where tone replaced a stem-final consonantal contrast can differ in whether a high tone evolved in stems that originally ended in a glottalic consonant or a low tone did, with the opposite tone evolving in stems ending in a non-glottalic consonant. Following Krauss (1979, this volume) and Leer (1979, 1999), languages that developed high tone from stem-final glottalic consonants are called 'high-marked' and languages that developed low tone

---

contrast shifted from the stem-final consonant to the preceding vowel just once.

[3] Some dialects of Koyukon are tonal (Hargus, p. c.), so the toneless dialects may have lost tone under the influence of neighboring non-tonal Western Alaskan languages after developing it from the stem-final glottal:non-glottalic contrast.

'low-marked'. The tone, high or low, is 'marked' because the opposite tone, low or high, developed on all other syllables: stems (and prefixes) that didn't end in a glottalic consonant.[4]

2.2 *The evolution of tone contrasts*

My account of tonogenesis in Athabaskan builds modestly on the proposals in Krauss (1979, this volume), and in Leer (1979, 1999, 2001). Krauss recounts how the accumulation of good descriptions of languages in this family gradually led Athabaskanists to treat tone as an innovation in some present-day daughter languages rather than an inheritance from the protolanguage. Leer lays out the conditions under which present-day tone evolved from non-tonal properties of the ancestral syllable rime, particularly in stems. (Because many more rime shapes are possible in verb and noun stems than prefixes, this discussion treats tonogenesis as a change in stems.)

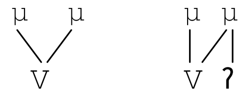
Fully five contrasts between the segments that compose rimes interact in determining how tone evolved, three between stem-final consonants and two between the preceding nuclei: (i) glottalic vs non-glottalic consonants (symbolized C' vs C), (ii) stops vs sonorants (K vs R), (iii) stops vs fricatives (K vs X), (iv) full vs reduced (or long vs short) nuclei, and (v) full vowel nuclei ending in a glottal stop vs those not ending with one. The glottalic consonants reconstructed stem-finally for the protolanguage by Krauss (1979), Leer (1979), and Cook & Rice (1989) include ejective stops and affricates, *t', *tɬ', *ts', *tʃ', tʃ'ʷ, k'ʲ, q', glottal stop *ʔ, and glottalized nasals, *m', *n', *ŋ', and glides, *w', *j', each of which, aside from *ʔ, has a non-glottalic counterpart. (Leer (1999) now reconstructs *tʃ'ʳ and *tʃʳ instead of earlier *tʃ'ʷ and *tʃ'ʷ.) Voiced and voiceless fricatives also contrast syllable-finally in stems, *l:*ɬ, *z:*s, *ʒ:*ʃ, *ɣʲ:*xʲ, *ʁ:*χ. (Rice (1997) collapses *j and *ʒ into a single proto-phoneme.) The preceding nuclei contrast in quantity: the stem vowel can be either full, i.e. long (VV), *i·, *e·, *a·, *u·, or reduced, i.e.

---

[4] The labeling of tones as 'marked' or not refers only to their diachronic origin and does not predict how they will behave synchronically; see also Rice & Hargus, this volume, for extensive discussion of what 'marked' means. For example, in Navajo (McDonough, 1999), the diachronically marked tone value is low but only a high tone spreads, from a disjunct prefix to the immediately following conjunct prefix. See Holton (this volume) for evidence that both high and low suffix tones are synchronically active in high-marked Tanacross.

short (V), *ə, *ɑ, *ʊ, where *ə is the reduced counterpart of both *i· and *e·.[5] Stems can end in no consonant at all if the vowel is full, *VV, or they can end in a glottal stop, *Vʔ, but stems can't end in just a reduced vowel, *V. As the vowel qualities of *Vʔ reflexes are (nearly) the same as those of *VV and differ in the same way as *VV reflexes from *V reflexes, *Vʔ must also consist of two moras, with the second coinciding with a ʔ as in (1). *VV are called 'modal' full vowels here, and *Vʔ are called 'glottalic' full vowels. The glottal stop is retained in reflexes in many tonal languages as well as in the non-tonal languages when it's not followed by a consonant. Like *VV nuclei, *Vʔ nuclei not only occur stem-finally but also before stem-final obstruents. As a result, *VʔK(') contrasts with *VVK('). However, VV does not contrast with Vʔ before sonorants.

(1) Contrast between *VV and *Vʔ rimes.



These five contrasts combine to produce the rime shapes for stems shown in Table 1.

	Stem-Final C						
Nucleus	None	Non-Glottalic C			Glottalic C'		
		Stop	Fricative	Sonorant	Stop	Fricative	Sonorant
VV	***VV***	***VVK***	***VVX***	***VVR***	***VVK'***	VV'X	VV'R
Vʔ	Vʔ	VʔK	VʔX	= VV'R	VʔK'	VʔX	= VV'R
V	(V)	***VK***	***VX***	***VR***	VK'	V'X	V'R

Table 1: *Rime shapes for Proto-Athabaskan stems.*

Marked tone evolves in stems whose rimes had the shapes represented in plain text in Table 1, unmarked tone in stems whose rime shapes are represented in bold

---

[5] In Athabaskanist practice, V· is used for full vowels and V for reduced ones. I follow that practice in citing actual examples, while using VV and V in formulae.

italics. The shape in the lower left cell is parenthesized because a stem's rime cannot consist of just a reduced vowel V; a prefix's rime can, however. The cells where the *Vʔ nucleus would occur before *R or *R' are occupied by = *VV'R because *Vʔ and *VV didn't, indeed couldn't, contrast before stem-final sonorants; see below for discussion. See also Table 2 below for the evolution of *VV'R and *V'R from etymological **VVR' and **VR'. Finally, *VʔX is the PA reflex of stems ending in a glottalic as well as a non-glottalic spirant. Shifting the glottal articulation to the preceding vowel is expected to produce *Vʔ'X, but this outcome obviously cannot differ from *VʔX.

The examples in (2) show how the evolution of tone depends on these five properties of the rime. The first four sets of examples factorially combine non-glottalic consonants (2a,c,f,h) vs glottalic consonants (2b,d,e,g,i), stops or affricates (2a-e) vs sonorants (2f-i), and (modal) full vowels (2c,d,e,h,i) vs reduced vowels (2a,b,f,g). The last two sets of examples show the contrasting evolution of stems whose nuclei consist of a modal full vowel (2j) vs glottalic full vowel (2k-n), either with no final consonant (2j,k) or a non-glottalic consonant (2l) vs glottalic one (2m,n). When the nucleus is a reduced vowel and the PA stem-final consonant is glottalic (2b,g), a high tone appears in the Chipewyan reflex and a low tone in the Gwich'in reflex, but when that PA consonant is instead non-glottalic (2a,f), the opposite tones appear, low in Chipewyan and high in Gwich'in. When the nucleus is a modal full vowel, high tone also appears in Chipewyan and low tone in Gwich'in when the PA final consonant is a glottalic sonorant (2i); cf. (2h) with a non-glottalic sonorant. However, when the nucleus is a modal full vowel and the final consonant is a glottalic stop or affricate, the opposite tones evolve in Chipewyan and Gwich'in (2d,e). A glottalic final consonant in the non-tonal language, Hupa, corresponds to high tone in Chipewyan and to low tone in Gwich'in, except when the PA consonant is a stop or affricate and the daughters show full vowel reflexes in their nuclei. The reflexes in (2d,e) show that the stem-final obstruent still loses its glottalic articulation in the tonal languages even though the marked tone did not evolve: the contrast between stems ending in glottalic vs non-glottalic stops and affricates merges completely when their vowel is full. Because a glottalic final consonant in Hupa otherwise corresponds to high tone in Chipewyan and low tone in Gwich'in, these latter languages are classified as 'high-marked' and 'low-marked', respectively. The set in (2k) shows that the marked tone also evolves in the tonal languages when the PA nucleus is a glottalic full vowel; compare the unmarked tones in (2j). A glottalic full vowel in fact always produces marked tone, regardless of whether the stem-final consonant is glottalic (2m,n) or not (2l).

(2) *Representative cognate sets. t', etc. = glottalic articulation, á, etc. = high tone, à, etc. = low tone, a·, etc. = full vowel, no · = reduced vowel, ą, etc. = vowel nasalization. (Examples from Krauss, 1979, this volume.)*

	Proto-Athabaskan	Chipewyan High	Gwich'in Low	Hupa Non-tonal
Stop, reduced vowel, VK vs VK' > unmarked vs marked tone				
a. "smoke"	**ɬəd*	*ɬə̀r*	*ɬád*	*ɬid*
b. "belly"	**wət'*	*bə́r*	*vàd*	*mət'*
Stop, full vowel, VVK and VVK' > unmarked tone				
c. "wife"	**ʔa·d*	*ʔà*	*ʔád*	*ʔad*
d. "scab"	**ɬu·t'*	*ɬùr*	*ɬɩ́d*	*ɬoh*
e. "spit"	**ʃʷe·q'*	*sèɣ*	*sʳɩ́g*	*se·ʔ*
Sonorant, reduced vowel, VR vs VR' > unmarked vs marked tone				
f. "bone"	**ts'ən*	*tθ'ə̀n*	*tθ'án*	*ts'iŋ*
g. "fire"	**qʊn'*	*kún*	*kǫ̀ʔ*	*xoŋ'*
Sonorant, full vowel, VVR vs VVR' > unmarked vs marked tone				
h. "rain"	**kʲa·n*	*tʃą̀*	*tsɩ́n*	*kʲaŋ*
i. "excrement"	**kʲʷa·n'*	*tsą́*	*trɩ̨̀ąʔ*	*tʃʷaŋ'*
No final C, modal vs glottalic full vowel, VV vs Vʔ > unmarked vs marked tone				
j. "stone"	**tse·*	*tθè*	*kɩ́·*	*tse·*
k. "foot"	**qeʔ*	*ké*	*kʷàjʔ*	*xeʔ*

	Proto-Athabaskan	Chipewyan High	Gwich’in Low	Hupa Non-tonal
Non-glottalic or glottalic stop, glottalic full vowel, VʔK, VʔK’ > marked tone				
l. “flour”	**ɬeʔdʒ*	*ɬéz*	*łùh*	*łi·ʍ*
m. “vein”	**kʲ’uʔts’*	*tʃ’ùdᶞé* [6]	*tʃ’ïǫ̀·*	*kʲ’its’*
n. “pitch”	**dʒeʔq’*	*dzé*	*dzìh*	*dze·χ*

Marked tone can also evolve on full vowel stems ending in a glottalic stop if the stop spirantizes, i.e. when VVK’ turns into VVX’. (Of course, only unmarked tone can evolve in stems ending in non-derived fricatives because they are never glottalic but only voiced or voiceless.) Leer (1979, 1999) proposes that spirantization and the consequent shift of glottalization to the preceding vowel are the crucial steps in the development of Proto-Athabaskan proper (PA, *forms) from an earlier, internally reconstructed stage of the protolanguage which he calls ‘Pre-Proto-Athabaskan’ (PPA, **forms). Spirantization and the shift of glottalization to the nucleus turns the lengthened allomorph **ʔa·tʃ’* of “few go” into **ʔa·’ʃ*, which evolves into marked tone *ʔás* in Chipewyan and *ʔò·* in Gwich’in, cf. the evolution of the perfective allomorph **ʔa·tʃ’-ŋ* into **ʔa·tʃ-ŋ* and ultimately unmarked tone *ʔàz* in Chipewyan and *ʔó·* in Gwich’in, when the final stop is blocked from spirantizing by the perfective suffix. See Leer’s papers for the morphological and phonological conditions on spirantization, and Leer (1979, 1999), Krauss & Leer (1981), Kari (1990), and Rice (1993, 1995) for the reconstruction of the perfective suffix. The stages in the evolution of the cognate sets in (2), including the effects of spirantization, are summarized in Table 2. Two PA entries are given for PPA full vowel stems ending in stops to show how spirantization affects the realization of a glottal articulation and what tone evolves subsequently.[7] Separate entries aren’t shown for reduced vowel stems because marked tone would evolve in them regardless of whether the final stop spirantized. Distinct forms are also proposed in some instances for the PA dialects that are ancestral to the tonal vs non-tonal daughters, thus revising Table 1.

---

[6] The low tone in the Chipewyan reflex *tʃ’ùdᶞé* of “vein” **kʲ’uʔts’* is unexpected, cf. the high tones in *ɬéz* “flour” and *dzé* “pitch”.

[7] Final stops spirantize under the same conditions in the non-tonal as the tonal languages, so the process entered the language before the protodialects ancestral to these two groups of languages diverged.

PPA	PA		Daughters	
	tonal	non-tonal	tonal	non-tonal
a. **VK	*VK	*VK	VK, unmarked	VK
b. **VK’	*V’K	*VK’	VK, marked	VK’
c. **VVK	*VVK	*VVK	VVK, unmarked	VVK
	*VVX	*VVX	VVX, unmarked	VVX
d. **VVK’	*VVK	*VVK	VVK, unmarked	VVK’
	*VV’X	*VVX	VVX, marked	VVX
f. **VR	*VR	*VR	VR, unmarked	VR
g. **VR’	*V’R	*VR’	VR(’), marked	VR’
h. **VVR	*VVR	*VVR	VVR, unmarked	VVR
i. **VVR’	*VV’R	*VVR’	VVR(’), marked	VVR’
j. **VV	*VV	*VV	VV, unmarked	VV
k. **Vʔ	*Vʔ	*Vʔ	VV(ʔ), marked	VVʔ
l. **VʔK	*VV’K	*VVK	VVK, marked	VVK
	*VV’X	*VVX	VVX, marked	VVX
m. **VʔK’	*VV’K	*VVK’	VVK, marked	VVK’
	*VV’X	*VVX	VVX, marked	VVX

Table 2: *Evolution of stems from PPA to PA to present-day daughters.*

In particular, stems ending in consonants with **Vʔ nuclei are represented with *VV’ for the ancestor of the tonal languages vs just *VV for the ancestor of the non-tonal ones. This captures the fact that the reflexes of these stems in the tonal languages have marked tone on a full vowel, while those in the non-tonal languages merely have a full vowel.

This table also shows the consequences of the two more general changes that distinguish tonal and non-tonal daughters: (i) in the tonal daughters, any glottal articulation that doesn’t shift off the stem-final consonant onto the preceding vowel is lost, and (ii) in the non-tonal daughters, stem vowels didn't phonologize constriction. Earlier, I suggested that these different evolutionary paths split PA into two distinct protodialects, which are represented separately in this table.

The smaller variety of reflexes of stems ending in sonorants in the daughter languages – the examples in (2g,h) exhaust the possibilities – shows that **VʔR doesn't contrast with **VVR' nor does **VʔR' contrast with **VVR'. This limitation was indicated in Table 1 by = VV'R in the cells that **VʔR and **VʔR' would have occupied. In sonorant-final stems with full vowel reflexes in the tonal daughters, marked tone always appears when the stems end in glottalic sonorants in the non-tonal daughters, and unmarked tone always appears when stems end in non-glottalic sonorants. (Only stems with full vowel reflexes need to be inspected because the reflexes of *Vʔ have the same vowel qualities as those of *VV.) Unlike stop-final stems, marked tone in a tonal language never corresponds to a non-glottalic sonorant in a non-tonal language. Therefore, a glottalic full vowel **Vʔ could have not occurred as a nucleus in stems ending in sonorants.

2.3 *The phonetics of glottal shift and subsequent tonogenesis*

Why, phonetically, does the evolution of marked tone depend on the manner of articulation of the stem-final consonant and on the fullness of the stem nucleus? Marked tone evolves when the glottalic articulation of the stem-final consonant shifts off that consonant onto the preceding nucleus, creating the same realization as when the nucleus itself was already glottalic. The first step in the phonetic shift from consonant to vowel is coarticulation of the vowel with the consonant's glottalic articulation. This coarticulation is an audibly different, constricted voice quality on vowels that preceded stem-final glottalic consonants from the modal voice quality occurring on those that did not (see §3 for detailed discussion). In the next step, the stem-final consonant lost its glottalic articulation, and the contrast had shifted to distinctive voice qualities on the preceding vowel. In the last step, the constricted voice quality was replaced by its characteristic F0, and the contrast became tonal. The vowels in **Vʔ nuclei can be treated as inherently constricted, as the second mora of these nuclei coincided with a ʔ, so voice quality in those nuclei was identical to those which became constricted by coarticulating with a stem-final glottalic consonant. The location of the contrast didn't really shift in stems with **Vʔ nuclei, only the constricted voice quality was later reinterpreted as tone in the same way.

The glottalic articulation can shift off the stem-final consonant if the preceding vowel already coarticulates sufficiently with that articulation that its voice quality is audibly constricted. What has therefore to be explained is how the consonant's manner of articulation and the vowel's fullness either permits or limits this coarticulation. Vowels became constricted and marked tone evolved before glottalic sonorants regardless of whether they were full or reduced. Furthermore, Vʔ nuclei didn't contrast with VV nuclei before sonorants. Full as well as reduced vowels also became constricted before fricatives derived from glottalic stops by spirantization. But only reduced vowels became constricted before glottalic stops that remained stops.

Stops differ acoustically from both sonorants and fricatives in that a burst of noise occurs when the oral closure is released and the air trapped behind it is permitted to escape. Kingston (1985, 1990) argued that laryngeal articulations, including the glottalic closure or constriction in ejective stops and affricates, must be coordinated with or 'bound' to the release in stops because those laryngeal articulations are conveyed to the listener by the ways in which they modify the noise burst and the onset of the transition to any following sonorant (see also Steriade, 1997). Because air flows continuously out of the mouth in both sonorants and fricatives, no acoustic event comparable to the stop burst occurs when their oral articulation is released, and therefore speakers aren't obligated to produce an accompanying laryngeal articulation at any time relative to the consonant's oral articulation.

Binding to the stop's release prevents the glottalic articulation from shifting off stem-final consonants with this manner, and thus correctly predicts that unmarked tone will evolve in syllables whose stem-final consonant remains a stop. (I postpone for a moment explaining why unmarked tone only evolves on full and not also reduced vowels before stem-final stops.) The absence of any acoustic event to bind to instead permits the glottalic articulation to shift freely to the preceding nucleus when the consonant is instead a sonorant or fricative.

Evidence reviewed in Kingston (1985; for instrumental evidence, see also Flemming, Ladefoged, & Thomason, 1994; Esling, Carlson, & Harris, 2002) shows that contrastive laryngeal articulations in postvocalic sonorants are often pronounced at the beginning of or before their oral constriction. If the glottalic articulation were timed in this way relative to the oral constriction in glottalic sonorants in PA, i.e. if */VR'/ were pronounced [V'R], then the glottalic articulation would already overlap with the preceding vowel. Not only would the vowel coarticulate enough with the sonorant's glottalic articulation for that articulation to shift readily to the syllable nucleus, but the pronunciation of /VVR'/ would be indistinguishable from that of /VʔR/ and /VʔR'/, and these sequences could not contrast.

Glottalic fricatives aren't reconstructed for PA. Crosslinguistically, they are in fact quite rare and little is known about their articulation. Maddieson, Smith, & Bessell (2001) describe ejective fricatives in Tlingit as being articulated with a long tight glottal closure in the middle of which the oral cavity is rapidly contracted, forcing air noisily out through the oral constriction. The acoustic consequences of these articulations are a period of silence before and after an interval of noise, and voice quality is often creaky in vowels abutting these periods of silence. The release of the glottal closure is often detectable as a burst or an abrupt onset of a following vowel. Because the glottis is closed well before the oral cavity is constricted and remains closed long after that constriction is released, the glottal closure always abuts a vowel; thus a glottalic /s’/ is pronounced [sʔ] in a syllable onset but [ʔs] in a syllable coda. The noise is ‘scrapey’ and often pulses, apparently because the oral constriction is narrow enough to open and close intermittently. The noise interval is not short, lasting 150 ms or more, but is consistently shorter than in non-glottalic voiceless fricatives in this language. No silent interval precedes or follows the noise interval of non-glottalic voiceless fricatives. If the glottalic articulation extended similarly far beyond the edges of the noise interval in the fricatives derived by spirantizing glottalic stops in PA, then it could constrict the voice quality of the preceding vowel and thereby easily shift off the consonant.

Structure preservation could also have contributed to this shift, as glottalic fricatives don’t occur lexically in PA. However, the glottalic articulation shifts as readily off glottalic sonorants which did occur lexically in PA, so structure preservation alone doesn’t motivate the shift. Instead the glottalic articulation shifts readily off consonants with both manners of articulation because it has no acoustic event to bind to in either of them.

If the glottalic articulation can’t shift off a stop because it’s bound articulatorily to its release, how did reduced vowels become constricted and then develop marked tone before glottalic stops, even while full vowels did not? For coarticulation to be interpreted as constriction, it must audibly alter the voice quality of enough of the preceding vowel that the vowel sounds different from one that isn’t followed by a glottalic consonant. If vowels always coarticulate to some fixed extent with the glottalic articulation of following consonants, regardless of their manner, then relatively more of a reduced than a full vowel will be constricted by this fixed amount of coarticulation. Ní Chasaide & Gobl (1993; Gobl & Ní Chasaide, 1999) show that in English and Swedish much or all of a vowel preceding an aspirated stop, i.e. one in which the glottis is spread open, is pronounced with a measurably breathy voice quality, which shows that the glottis opened long before the oral closure is made. The effect on the preceding vowel is similar to but less extreme than the preaspiration observed in Icelandic (Thrainsson, 1978; Kingston, 1990). If a constricted voice quality extends equally far into a vowel before a glottalic consonant, a reduced vowel’s voice quality could thus differ more before glottalic vs non-glottalic consonants than would a full vowel’s in the same contexts, most of which would still be pronounced with modal voice in both contexts. A reduced vowel could thereby become constricted enough before a glottalic stop via coarticulation alone for marked tone to evolve even without the laryngeal articulation shifting onto the vowel.

This scenario is explicitly avoided in present-day Hupa, where a syllable-final glottalic consonant's laryngeal articulation is realized as creaky voice during the second half of a preceding full vowel, but on the consonant itself when the preceding vowel is reduced (Gordon, 2001). Gordon argues that this difference arises from a constraint against making the voice quality of the entire syllable nucleus creaky or non-modal: a reduced vowel would be creaky from beginning to end if the consonant's laryngeal articulation were pronounced in the same way as it is when a full vowel precedes. The need for such a constraint in Hupa shows that it was probably easy in PA for much or all of a reduced vowel to become constricted merely by coarticulating with a following glottalic consonant, even in a stop where the laryngeal articulation was bound to the release.

Golla's (1976, p.c.) description of the phonetic realization of glottalic consonants stem-finally in the non-tonal Pacific Coast language Tututni presents a similar apparent problem: the glottalic articulation is realized before the oral closure in stem-final stops, but after it in stem-final sonorants. This timing would suggest that glottalization is more likely to shift to the preceding vowel when the final consonant is a stop than a sonorant because it's already closer to the vowel. Here, too, I would suggest instead that the laryngeal articulation is timed relative to the oral one in a way that ensures the laryngeal articulation remains a property of the stem-final consonant. Because the laryngeal articulation is bound to the release in stops it can begin earlier without shifting off them, but because it has nothing to bind to in sonorants, it must occur later to avoid shifting to the preceding vowel.

2.4 *Summary*

In this section of the article, I have argued that the present-day distribution of tone in Athabaskan languages arose from the shift of the laryngeal articulation of a stem-final glottalic consonant to the preceding vowel in the protodialect that's ancestral to the Canadian, northeastern Alaskan, and Apachean languages but not in the protodialect ancestral to the Pacific Coast or southwestern Alaskan languages. Speakers of the pre-tonal protodialect migrated from the homeland east and south into Canada and then down the eastern face of the Rockies, eventually reaching what is now the southwestern United States. Other speakers of the pre-tonal protodialect migrated north and west into Alaska in the wake of some speakers of the non-tonal protodialect. Other speakers of the non-tonal protodialect migrated down the Pacific coast early on, eventually reaching present-day California. The argument's keystone is that glottalic obstruents no longer contrast with non-glottalic consonants stem-finally in the tonal languages, while they still do in most of the non-tonal languages – all of the Pacific Coast languages and all but Deg Hit'an, Holikachuk, and Koyukon in western Alaska.

The grammatical distribution of the marked tone that evolved from vowel constriction within the tonal languages was shown to depend on the fullness and voice quality of the stem vowel and the manner of articulation of the stem-final consonant. Reduced stem vowels always bear marked tone when the stem ended in a glottalic consonant, but full vowels only do when the stem-final glottalic consonant was a sonorant or fricative at the time vowels became constricted, and not when it was a stop or affricate. Spirantization of some stops at the end of full vowel stems freed a glottal articulation to shift to the preceding vowel in the ancestor of the tonal languages, and marked tone evolved even though the stem vowel was full. The evolution of marked tone in stems whose final consonant was not glottalic was accounted for by reconstructing nuclei of the shape **/Vʔ/ in these stems.

Unmarked tone evolved in full vowel stems ending in a glottalic stop because the glottalic articulation was bound articulatorily to the noise burst at the stop's release and most of the preceding vowel remained modal voiced even if the vowel coarticulated to a fixed extent with the stop's glottalic articulation. A fixed amount of coarticulation was apparently enough to constrict reduced vowels audibly, and they instead acquired marked tone before glottalic stops. Full as well as reduced vowels became constricted and acquired marked tone before glottalic sonorants and fricatives because there was no acoustic event binding their laryngeal articulation, and it was likely to have been pronounced before the consonant's oral articulation. Constriction would therefore have been likely to extend further back into preceding vowels, enough to audibly constrict the voice quality of full as well as reduced vowels. The realization of the glottalic articulation before the oral one in sonorants would also have made it impossible to produce a contrast between **/VʔR/, **/VʔR'/, and */VVR'/ sequences, as all three would have been pronounced [VV'R].

I next turn back to the first question of this article: how can a low tone have evolved from stem-final glottalic consonants in some Athabaskan languages while a high tone has evolved from this source in others? §§3-4 deal in turn with the two phonetic answers to this question: how opposite tones evolved early vs late in the history of the tonal Athabaskan languages.

3. *A phonetic explanation for the early evolution of high and low tone in Athabaskan*

3.1 *The articulation of glottalic consonants and the evolution of constricted vowels*

In this section of the article, I describe how the glottis is constricted in glottalic consonants and how that constriction could have shifted to the preceding vowel. I also present evidence that languages differ in how their speakers pronounce these consonants in ways that could have led first to the development of either creaky or tense voice qualities and ultimately to the development of either low or high tone.

3.1.1 *Glottalic stops, a.k.a. ejectives.* According to standard descriptions of ejective stops and affricates, the glottis is closed tightly to isolate the oral cavity aerodynamically from the subglottal cavity, and the oral cavity is then maximally contracted to compress the air inside it. Compression produces an extremely intense burst when the oral closure is finally released, and voicing only begins following a long delay because the glottis remains tightly closed for long after the release. Once voicing begins, F0 is high and the voice quality is modal or even tense, the latter if the speaker continues to compress one vocal fold against the other medially. Ejectives are described as being pronounced in this or a similar way in Tigrinya (Fre Woldu, 1985; Kingston, 1982, 1985), Montana Salish (Flemming, Ladefoged, & Thomason, 1994), Tsez (Maddieson, Rajabov, & Sonnenschein, 1996), Chipewyan (Hogan, 1976), Tlingit (Maddieson, Smith, & Bessell, 2001), Navajo (Lindau, 1984, McDonough & Ladefoged, 1993), Western Apache (Gordon, Potter, Dawson, de Reuse, & Ladefoged, 2001), Hupa (Gordon, 1996), and perhaps Minto (Lower Tanana, Tuttle, 1997, 1998). Ejectives can also be pronounced with glottal closure or constriction but without either extreme contraction of the oral cavity and compression of the air inside it or the glottis remaining tightly closed for a long time after the release. In this pronunciation, the burst is not dramatically more intense than in a pulmonic voiceless stop, voicing begins soon after the release, and at voice onset, F0 is low and voice quality is creaky. This pronunciation or one similar to it has also been observed in Tigrinya (Kingston, 1985), apparently when Tigrinya speakers don't hyper-articulate these consonants, Quiché (Kingston, 1982; Pinkerton, 1986), Hausa (Lindau, 1984), Hadza (Sands, Maddieson, & Ladefoged, 1993), Gitskan (Ingram & Rigsby, 1987), Witsuwit'en (a.k.a. Babine; Wright, Hargus, & Davis, 2002), and as an alternative pronunciation in Minto (Tuttle, 1997, 1998). Bird (2002) reports that in the Athabaskan language Dakelh (Carrier), the same speaker may pronounce ejectives in either way, although typically the ejective affricates /ts’, tʃ’/ and secondarily articulated stops /k^{w}’, t^{l}’/ are pronounced the first way and the unadorned ejective stops /t’, k’/ the second.[8] (/tˡ’/ is a laterally released stop followed by a lateral consonant, approximately [tˡʔl], not a lateral affricate, in this language.) I will refer to these two pronunciations from now on as ‘stiff’ and ‘slack’ ejectives, where ‘stiff’ and ‘slack’ refer to the state of the vocal folds during the interval when the glottis is constricted.[9] If F0 and voice quality differ similarly in vowels preceding these two pronunciations of ejectives as they do in vowels

---

[8] These examples have all been analyzed instrumentally, in one way or another. Impressionistic descriptions of the pronunciation of ejectives in other languages indicate that this crosslinguistic variation is in fact wide-spread. For example, Sapir (1922) describes the ejectives — which he called “fortes” — in Takelma as “pronounced with the characteristic snatched or crackly effect (more or less decided stress of articulation of voiceless stop followed by explosion and momentary hiatus) prevalent on the Pacific Coast” (p. 33). He goes on to describe how in the Takelma ejectives, “the glottis is closed just before or simultaneously with the moment of consonant contact, is held closed during the full extent of the consonant articulation, and is not opened until *after* the consonant release...The cracked effect of the fortes ... is due to the sudden opening of the closed chamber formed between the closed glottis and the point of consonant contact...; the hiatus generally heard between a fortis and following vowel is simply the interval of time elapsing between the consonant release and the release of the glottal closure.” (pp. 33-34, his italics). In contrast, Boas (1911) describes the pronunciation of ejectives in Nass River Tsimshian as using “the hiatus [glottal closure: JK] frequently, without, however, giving the preceding stop enough strength to justify the introduction of a fortis.” (p. 288). According to Tarpent (1987), Tsimshian’s “glottalized stops are preglottalized” and “glottalization is not very strong and is at times barely perceptible.” (p. 37).

[9] In my earlier ms., I referred to these two pronunciations as ‘tense’ and ‘lax’ instead. ‘Stiff’ and ‘slack’ capture better the difference in the state of the folds that would produce tense vs creaky voice and ultimately high vs low tone, respectively. Halle & Stevens (1971; see also Stevens, 1977) introduced these terms to describe the state of the vocal folds in stops contrasting for [voice]. They propose that the folds are stiff in [-voice] stops to inhibit voicing and slack in [+voice] stops to encourage it. Fold stiffening cooperates with pulling the vocal folds apart to prevent voicing, and fold slackening cooperates with bringing them together to produce it. Halle & Stevens (1971), Stevens (1977), and also Löfqvist, Baer, McGarr, & Seider Story (1989) propose that cricothyroid contraction stretches and thus stiffens the folds in [-voice] stops by tilting and sliding the thyroid cartilage forward with respect to the cricoid cartilage. A side effect of the stiff/slack difference is higher F0 in vowels flanking [-voice] than [+voice] stops. Here, I adopt the proposal that F0 is determined by whether the cricothyroid muscles are contracted (see §3.2), but I propose that the folds may be stiffened or slackened by this means independently of other laryngeal articulations, specifically when the vocal folds are constricted.

following them, then a stiff pronunciation by speakers of one protodialect could have been the direct phonetic source of marked high tone, and a slack pronunciation by speakers of another protodialect the direct phonetic source of marked low tone.[10] Just what 'stiff' and 'slack' pronunciations are is described in §3.2 below.

3.1.2 *Glottalic sonorants*. That marked tone always develops before a stem-final glottalic sonorant in the tonal Athabaskan languages shows that the glottal closure's effects on voice quality and F0 were a reliably perceived phonetic property of this manner of articulation. Because air flow out of the mouth or nose is unimpeded in a sonorant, air cannot be trapped and compressed between the glottal and oral closures in a glottalic sonorant. The goal of the glottal closure in sonorants must be acoustic rather than aerodynamic. The acoustic effects of the glottal closure are to interrupt the signal briefly, or if the glottis is only constricted and not firmly closed and the folds can vibrate, to produce a non-modal voice quality.[11] (Because oral air pressure cannot rise much in a glottalic sonorant, voicing is also more likely than in a glottalic stop even if the glottis is tightly constricted.)

Speakers are free to produce this non-modal voice quality as tense or creaky voice, i.e. in such a way as to raise or lower F0. Furthermore, speakers' choices about which non-modal voice quality to use in pronouncing glottal sonorants probably determined how they chose to pronounce the glottalic stops and affricates, ensuring that the same tone developed from glottalic consonants in both manners of articulation. Solnit & Kingston (1988; Kingston & Solnit, 1989) show that Southeast and East Asian languages also differ in whether higher or lower tonal reflexes develop after initial glottalized, aspirated, and voiced consonants in tone splits. These differences probably also depend on sonorants' freely permitting speakers to produce these articulations with either stiff or slack vocal folds in sonorants, which are then extend to stops.

---

[10] Striking evidence that a glottalic articulation can be pronounced so as to raise or lower F0 is found in Carrier (Dakelh), where syllable-initial glottal stops lower pitch on following vowels, but syllable-final glottal stops raise pitch on preceding vowels (Pike, 1986; Story, 1989).

[11] The glottal closure or constriction is often described as occurring at the beginning of or even before the oral articulation of a glottalic sonorant: "the glottal closure is synchronous with the momentary voiceless initial phase of the continuants [= sonorants, JK], its release being immediately followed by the voiced phase of the continuant" (Sapir, 1938, p. 249), increasing its likelihood of affecting voice quality and F0 in preceding vowels. For instrumental evidence of this timing pattern see Flemming, et al. (1994) and Esling, et al. (2002).

3.1.3 *Changing the voice quality of the preceding vowel.* These differences between languages in how their speakers pronounce the glottal closure or constriction in consonants could lead directly to the evolution first of tense vs creaky voice quality and ultimately of high vs low tone. For these stiffness differences to have been the phonetic source of high and low tone on vowels preceding glottalic consonants in the Athabaskan languages, they must have arisen before as well as after the oral constriction of the consonant. This is likely in glottalic sonorants because the laryngeal articulation precedes the oral one. If, on the other hand, the laryngeal articulation is bound to the release in stops, how can it affect the preceding reduced vowel enough to constrict it? The breathy articulation of vowels preceding English and Swedish voiceless aspirated stops (Ní Chasaide & Gobl, 1993; Gobl & Ní Chasaide, 1999) and Icelandic preaspirated stops (Kingston, 1990) shows that the preceding vowel's voice quality can be affected even if the consonant's laryngeal articulation is bound to its release. If laryngeal articulations of all kinds are anticipated in this way, then a vowel would also be measurably constricted before a glottalic stop. For this constriction to replace the glottalic articulation of the stop as the bearer of the contrast, it would have had to become audible, as a distinct non-modal voice quality on that vowel. The timing of the laryngeal articulation before the oral one in sonorants did not, however, lead speakers to adopt the same relative timing of the two articulations in stops, too. Even though the glottalic articulation of sonorants shifted to the preceding vowel, speakers didn't also shift the glottalic articulation of stops, too. The glottalic articulation instead remained bound to the stop's release. Whatever the extent of laryngeal coarticulation between the vowel and the following stop, a distinctly constricted voice quality therefore was only audible enough on reduced vowels before glottalic stops to eventually replace the consonant's glottalic articulation as the contrast bearer.

3.2 *The physiology of glottal constriction and F0*

3.2.1 *Functional anatomy of the larynx.* The larynx is built on three cartilages and a bone (Figure 1, adapted from jimswan.com/anatomy/respiration_notes.htm.). Its base is the signet ring-shaped cricoid cartilage sitting on top of the trachea. The thyroid cartilage is two vertically oriented sheets or laminae of cartilage that meet at an angle at the front. The left lamina's inner surface is shown in the top panel of Figure 1, in rest position and also tilted and slid forward (dotted outline). Protruding up and down from each lamina's back edge are the superior and inferior horns; the superior horns attach to the posterior tips of the U-shaped hyoid bone above (not shown), and the inferior horns articulate with facets on the sides of the cricoid cartilage below; this articulation is shown by the dotted outline of the left inferior horn of the thyroid cartilage in the top panel. Finally, sitting on cylindrical facets atop either side of the cricoid cartilage's posterior bulge are the tetrahedral arytenoid cartilages; the left arytenoid cartilage is shown in the top panel. Each vocal fold attaches posteriorly to the forward pointing vocal process on each arytenoid cartilage, and runs forward to attach anteriorly at the juncture of the thyroid laminae, as can be seen in the middle and bottom panels.

Speakers can (i) close or open the glottis, (ii) stretch or slacken the vocal folds, and (iii) stiffen the folds and press them medially against one another or relax them and allow them not make complete contact all along their length. First, contracting the interarytenoid muscles (INT, bottom leftl), which connect the backward pointing muscular processes of the arytenoid cartilages, rocks them toward the midline and brings the two vocal folds together, closing the glottis (adducting). Closing the glottis is probably aided by contracting the lateral cricoarytenoid muscles (LCA, middle right), too, as lateral cricoarytenoid contraction pulls forward on the muscular processes, sliding the arytenoid cartilages forward on their facets and bunching the folds. Contracting the posterior cricoarytenoid muscles (PCA, middle left) instead pulls laterally on the arytenoid cartilages' muscular processes, rocks them away from the midline, pulls the vocal folds apart, and opens the glottis (abducting). Second, contracting the cricothyroid muscles (top) rocks and slides the thyroid cartilage forward with respect to the cricoid cartilage. These motions pull on and stretch the vocal folds because they are attached to the thyroid cartilage's inner surface. Finally, contracting the portion of the thyroarytenoid muscles lying just lateral to the glottis, the thyrovocalis (TA bottom right), shortens the folds and causes them to bunch medially and to press one against the other if they've been adducted.[12]

---

[12] The thyroarytenoid muscles can be divided into medial and lateral parts, called the thyrovocalis and thyromuscularis, respectively. Here, I will refer to the thyrovocalis as the thyroarytenoid, because it is the contraction of this medial part of the muscle that changes the stiffness of vocal folds and the degree of their medial compression against one another.

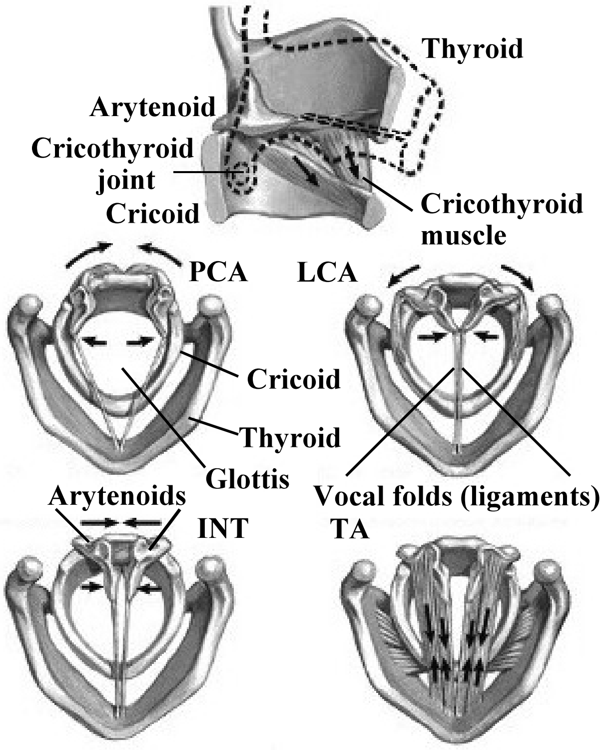


Figure 1: *The laryngeal cartilages and muscles: (top) right lateral view of inner surface of the left half of the laryngeal cartilages – right is front – (middle and bottom) top views looking down into the glottis – bottom is front.*

3.2.2 *Glottal constriction in consonants.* Instrumental studies of glottalic consonants themselves have analyzed only their acoustics or aerodynamics, and not the activity in the muscles responsible for producing the glottal closure or constriction. We must instead look to electromyographic investigations of the tense stops in Korean, of Danish *stød*, and of glottal stop. Hirose, Lee , & Ushijima (1974) and Hirose, Park, Yoshioka, Sawashima, & Umeda (1981) observed that the thyroarytenoid and lateral cricoarytenoid muscles are more active in the tense than the lax stops in Korean during the last few tens of milliseconds before the release of the oral closure. The contraction of these muscles closes the glottis, tightens the glottal closure by increasing the force with which one vocal fold presses medially against the other, and stiffens the folds. Medial compression and stiffening also prevent the folds from vibrating until the tense stop is released, despite the fact that they are adducted (Kagaya, 1974). The thyroarytenoid muscles also contract in the production of both Danish *stød* (Fischer-Jørgensen & Hirose, 1974) and glottal stops (Hirose & Gay, 1972), and in glottal stops the lateral cricoarytenoid muscles do, too (Hirose & Gay, 1972). These electromyographic data all show that the thyroarytenoid and perhaps also the lateral cricoarytenoid muscles contract when the folds are pressed firmly against one another to close or constrict the glottis.

The stiffening of the folds in the Korean tense stops also raises F0, produces a faster rise in intensity, and excites higher frequency resonances more at the beginning of the following vowel than after a lax stop, all correlates of a tense voice quality (Han & Weitzman, 1970; Hardcastle, 1973; Kagaya, 1974; Hirose, Park, Yoshioka, Sawashima, & Umeda, 1981). If Proto-Athabaskan speakers contracted the thyroarytenoid muscles to close the glottis before their glottalic consonants as Korean speakers do after their tense stops, a tense voice quality and ultimately a high tone could have evolved directly. How then can some of the tonal Athabaskan languages have evolved in the opposite direction, to creaky voice and ultimately a low tone, if closing the glottis tightly by thyroarytenoid contraction stiffens the folds? Stiffness of the vocal folds and thus both tense vs creaky voice quality and high vs low F0 are all determined by two independent mechanisms, contracting the cricothyroid or the thyroarytenoid muscles. The anatomy of the vocal folds themselves and their response to forces acting on them permits the speaker to control independently the force with which one fold is pressed against the other and the effective stiffness of the folds, such that the glottis can be closed tightly without stiffening the vibrating portion of the folds. Glottal closure can therefore be as direct a phonetic source of creaky voice and low tone as tense voice and high tone.

The vocal folds themselves can be divided anatomically into a cover and body, whose stiffness can be independently controlled, by contraction of the cricothyroid and thyroarytenoid muscles, respectively. This independent control permits the speaker to vibrate the vocal folds at different rates while maintaining the same glottal constriction. In §3.2.3, I first describe the anatomical division of the vocal folds into a cover and body and the independent effects of contracting the cricothyroid or thyroarytenoid muscles on cover and body stiffness. I then discuss the dynamics of F0 and voice quality control informally. The discussion turns next to the separate innervation of the cricothyroid and thyroarytenoid muscles and how the extreme voice qualities, falsetto and creaky voice, may be produced by strongly contracting one of these muscles alone, and finishes by showing how tense voice is produced by contracting both muscles at once. The following section sketches the historical stages of evolution of creaky and tense voice into low and high tone, respectively, in the tonal Athabaskan languages (§3.2.4).

To anticipate the message of the rest of §3.2, strong cricothyroid contraction without any thyroarytenoid contraction brings only the upper margins of the folds into contact, lengthens the open phase of the folds' vibratory cycle, limits vibration to the fold's cover, and raises F0 extremely. The resulting voice quality is falsetto. Strong thyroarytenoid contraction without any cricothyroid contraction brings the lower as well as the upper margins into contact, increases medial compression of the folds, lengthens the closed phase of the vibratory cycle, also limits vibration to the cover, and lowers F0 extremely, if it is strong enough to prevent vibration from extending into the body of the folds. The resulting voice quality is creaky voice. This voice quality is expected in vowels flanking a glottalic consonant as the speaker begins or stops producing the glottal closure or constriction characteristic of these consonants. F0 is expected to be very low in these vowels. This voice quality and its low F0 is thus the one expected next to stops pronounced with the glottis extremely constricted, not tense voice and its high F0. A tense voice quality and high tone can nonetheless evolve as directly from a glottalic consonant as a creaky voice quality and low tone in one of two ways, because speakers may also contract the cricothyroid muscles at the same time as the thyroarytenoid muscles. That is, they may produce the glottal closure or constriction in a glottalic consonant with the cover as well as the body stiff, and the resulting tense voice quality and high F0 can also be transferred to vowels that coarticulate with these consonants. Alternatively, the thyroarytenoid muscles may not be contracted enough to prevent vibration from extending into the body of the folds, and as a result F0 may be raised. Moderate but not extremely strong thyroarytenoid contraction is probably the source of raised F0 next to tense stops in Korean, because the cricothyroid muscle isn't especially contracted when they are pronounced. High tone is more likely to evolve in Athabaskan when the cricothyroid as well as the thyroarytenoid muscles are contracted rather than through moderately strong thyroarytenoid contraction alone because extreme thyroarytenoid contraction is required to close the glottis tightly. The discussion in §3.2.4 focuses on this possibility.

3.2.3 *Stretching the folds' covers, stiffening their bodies, or both.* The key to understanding how F0 and voice quality are determined by the state of the vocal fold tissue is that the folds consist of a relatively inflexible body surrounded by a relatively flexible cover, whose stiffnesses can be independently changed (Hirano, 1974, 1975, 1977; Hirano, Kurita, & Nakashima, 1981; Hirano & Kakita, 1985; Alipour-Haghighi & Titze, 1991; Cooper, Partridge, & Alipour-Haghighi, 1993;

Titze, 1994, Story & Titze, 1995). Stretching the cover or contracting the body determines the rate of vocal fold vibration by changing the fold's length and stiffness.

How then are the length and stiffness of the vocal folds changed and how do those changes determine F0? The vocal folds are lengthened by contracting the cricothyroid muscles, which tilts and slides the thyroid cartilage forward with respect to the cricoid cartilage (Figure 1, top panel). Tilting and sliding the thyroid cartilage forward pulls on the vocal folds attached to its inner surface, lengthening them. Lengthening the folds stretches the cover and stiffens it, and may also stretch the body of the folds if the thyroarytenoid muscles aren't contracted too strongly. Contracting the thyroarytenoid muscles instead shortens the folds and thereby reduces the stiffness of the cover, and if the glottis is adducted, compresses one fold against the other (Figure 1, bottom right panel). These antagonistic changes in cover length, stretch, and stiffness should raise and lower F0, respectively. However, contracting the thyroarytenoid muscles also stiffens the folds' bodies, so whether F0 is actually lowered by thyroarytenoid contraction depends on whether the contraction of these muscles permits vibration to extend into the folds' bodies. If vibration extends no more deeply than the slackened cover, then contracting the thyroarytenoid muscles lowers F0 but if vibration extends deeper, into the body itself, then contracting these muscles raises F0 instead. Stiffness and thus F0 correlate directly with both the stress applied to the folds by stretching them and the effective mass of the vibration (Titze, 1994; Story & Titze, 1995). Stress itself increases with the strain imposed by stretching the folds. Alipour-Haghighi & Titze (1991; also Cooper, Partridge, & Alipour-Haghighi, 1993) measured the relationship between stress and strain separately in the vocal fold cover and body and found that stress increases in the cover much more linearly with increasing strain than it does in the body. Even more important, across a wide range of relatively low strain values (length increases of 0-20%), stress increases in the body are small, and up to strains of about 35% they are smaller than in the cover. Because increases in vocal fold length across this large range don't change body stress as much as cover stress, F0 will be determined more by the stress on the cover than on the body for most length increases brought about by contracting the cricothyroid muscles. Cricothyroid contraction is therefore predicted to stretch the folds' covers and increase their rate of vibration across a very large range more or less independently of the stiffness of their bodies.

A speaker can contract the cricothyroid muscles without contracting the thyroarytenoid muscles and vice versa because the muscles are separately innervated, the cricothyroid muscles by the superior laryngeal branch of the vagus nerve and the thyroarytenoid muscles by its recurrent laryngeal branch. Rubin

(1963) and Hirano (1974) showed that stimulating the superior laryngeal nerve lengthened and thinned the vocal folds and increased their rate of vibration. Koike, Hirano, Morio (1974), Hirano (1974), and Hirano, Matsuo, Kakita, Kawasaki, & Kurita (1983) report that stimulation of the thyroarytenoid muscles through the recurrent laryngeal nerve causes the vocal folds instead to shorten, thicken, and bulge toward the midline, all of which would tighten the constriction of the glottis if the folds were adducted.

In line with Alipour-Haghighi & Titze's (1991) finding that increasing strain increases cover stress more linearly than body stress, Moore & Berke (1988) found that independently stimulating the superior laryngeal nerve raised F0 dramatically, up to as high as 340 Hz, but independently stimulating the recurrent laryngeal nerve did not raise F0 nearly as much, to only 190 Hz. Voice quality was also affected in diametrically opposed ways by stimulation of the two nerves. With increasing superior laryngeal nerve stimulation and cricothyroid contraction, the open portion of each glottal cycle increased, subglottal air pressure increased only slightly, lateral movement of the folds decreased, and the lower margins of the folds eventually stopped making contact. When only the upper margin of the folds made contact, the voice quality shifted from modal to falsetto. With increasing recurrent laryngeal nerve stimulation, on the other hand, the closed portion of the glottal cycle and subglottal air pressure both increased, the length of the folds' vibrating portion decreased, and the lower as well as the upper margins of the folds made increasingly firmer contact. The thyroarytenoid contraction caused by stimulating the recurrent laryngeal nerve shifted voice quality from modal to tense voice rather than to falsetto voice. Although these authors don't directly report data on combined stimulation of both nerves, they nonetheless assert that "[recurrent laryngeal nerve stimulation] and [superior laryngeal nerve stimulation] demonstrated the [cricothyroid] muscles' ability to thin and lengthen the folds for any given level of [recurrent laryngeal nerve stimulation]" (p. 719). This is possible if cover stiffness can be controlled independently of the body stiffness, as their separate innervation suggests they can be.

Hirano (1974) and Story & Titze (1995) propose that vibration extends no more deeply into the fold than the cover in falsetto voice because the vocal ligament at the boundary between the cover and body takes up the longitudinal tension applied to the fold rather than the thyroarytenoid muscle itself, even though that muscle may be lengthened considerably by strong cricothyroid contraction. The absence of any active increase in body stiffness due to the thyroarytenoid's relaxation also keeps the fold's lower margin away from the glottal midline throughout the vibratory cycle in this voice quality. Other studies comparing muscle activity levels during the production of falsetto vs chest register confirm the observation that F0 is controlled by cricothyroid contraction alone in falsetto register but jointly by cricothyroid and thyroarytenoid contraction in chest register (Hirano, Ohala, & Vennard, 1969; Hirano, Vennard, & Ohala, 1970; Shipp & McGlone, 1971; Baer, Gay, & Niimi, 1976; Titze, 1994).

At the other extreme of the speaker's F0 range, where voice quality shifts from modal to creaky voice (a.k.a. 'vocal fry' or 'pulse register'), the physiological evidence shows that it is the thyroarytenoid muscles that are strongly contracted and the cricothyroid muscles that are relaxed (McGlone & Shipp, 1971) . As a result, one vocal fold presses medially against the other, which lengthens the closed phase of the vibratory cycle, decreases the amount of air that passes up through the glottis each cycle, and often increases subglottal air pressure by trapping more air below the glottis (McGlone, 1967; McGlone & Shipp, 1971; Murry, 1971; Murry & Brown, 1971; Whitehead, Metz, & Whitehead, 1984; Moore & Berke, 1988; Childers & Lee, 1991; Blomgren, Chen, Ng, & Gilbert, 1998). Allen & Hollein (1973), Hollein (1974), and Zemlin (1988) all observed that in creaky voice the vocal folds are also thicker and shorter and that only the anterior portion of the cover vibrates. Allen & Hollein showed, too, that the lower surface of the ventricular or 'false' folds may be pressed down on the upper surface of the 'true' vocal folds during creaky voice, impeding body vibration by increasing the folds' inertia. Because the vibration is limited to the cover, the effective mass of the folds is smaller. The absence of cover stiffening, the thickening and shortening of the folds, and the smaller effective mass all reduce the rate of vocal fold vibration markedly, to as little as a third the F0 of an adult male's low pitched modal voice (McGlone, 1967; McGlone & Shipp, 1971; Blomgren, et al., 1998).

Stiffness and F0 aren't only proportional to how stretched the cover is, but also to the mass of vibrating tissue, i.e. to how deeply the vibration goes into the tissue. Although F0 may be very low when only the thyroarytenoid muscles are contracted, much of the data on the relationship between F0 and thyroarytenoid contraction show that F0 also rises when these muscles are contracted, at least in modal voice or chest register. Moore & Berke (1988) observed F0 raising with recurrent laryngeal nerve stimulation, which caused the thyroarytenoid muscles to contract. The studies of F0 variation while singing in chest register also showed that thyroarytenoid activity increased with F0 (Hirano, Ohala, & Vennard, 1969; Hirano, Vennard, & Ohala, 1970; Shipp & McGlone, 1971; Baer, Gay, & Niimi, 1976; Titze, 1994), as do studies of F0 and thryoarytenoid activity in speaking (Ohala, 1970; Atkinson, 1978). Titze (1994) even found that F0 was determined more by thyroarytenoid than cricothyroid contraction at the lower end of a singer's range.

The resolution of this apparent paradox falls out of the physiological differences between modal and creaky voice. First, thyroarytenoid muscle contraction is not extreme in modal voice, so the vibration extends into the bodies of the folds, as well as stiffening them. Increasing the mass of vibrating tissue and stiffening the body increases F0. Moreover, when increased thyroarytenoid contraction raises F0 in modal voice, the cricothyroid muscles are also active, even if their effect on F0 is not as strong as the thyroarytenoid muscles' at the low end of the speaker's modal range. Increasing the strain applied to the folds' covers by contracting the cricothyroid muscles also increases F0. That is, both muscles' contraction determines the effective stiffness of the folds – thyroarytenoid contraction stiffens their bodies and cricothyroid contraction stiffens their covers – and their bodies vibrate as well as their covers.

On the other hand, in creaky voice, like falsetto, vibration extends no deeper than the cover of the folds. Vibration extends no deeper than the cover in this voice quality because the thyroarytenoid muscles are contracted so much that the body simply cannot vibrate. Alipour-Haghighi & Titze (1985, 1991; Alipour-Haghighi, Titze, & Perlman, 1989) show that when the thyroarytenoid muscles are contracted as strongly as possible, stress is easily three times as large as that applied to the vocal fold body by maximal cricothyroid contraction. According to Titze, Luschei, & Hirano (1989), voicing should either be inhibited or highly 'pressed' under these circumstances so long as cricothyroid relaxation leaves the cover unstressed. Even if the cricothyroid muscles are contracted, stretching the folds adds little to the stress on the body because this extreme thyroarytenoid contraction stresses the body so much. Furthermore, extreme thyroarytenoid contraction prevents vibration from extending into the body of the folds in creaky voice, and the rate of vibration is therefore independent of body stress. If the speaker does nothing to stretch the cover, F0 will of course be low under these extreme circumstances. However, even such extreme thyroarytenoid contraction in no way prevents the speaker from also contracting the cricothyroid muscles, turning creaky into tense voice and raising F0. Tense voice resembles falsetto voice in that only the cover vibrates, at a high rate, while at the same time resembling creaky voice in that the extreme thyroarytenoid contraction brings the lower as well as the upper edges of the folds into contact.

3.2.4 *Getting to low and high tone from creaky and tense voice.* The first stage in the development of tone from stem-final glottalic consonants is speakers' producing either creaky or tense voice in vowels preceding glottalic consonants. At this stage, this non-modal, constricted voice quality results simply from coarticulation between the vowel and consonant. The initial portion of *Vʔ nuclei would also have been pronounced with this voice quality. Whether the constricted voice quality is creaky or tense depends on how the glottalic consonants and the ʔ in *Vʔ are articulated. If the vocal folds' covers are slack in these articulations, the constricted voice quality will be creaky but if they are stiff, it will instead be tense. In the next stage, the glottal closure or constriction is no longer made during the consonant itself and the burden of the contrast between glottalic and non-glottalic consonants shifts to the difference between the constricted vs modal voice quality on the preceding vowel. The final stage is a shift from speakers' producing one or the other constricted voice quality with its characteristic F0 value, low in creaky voice and high in tense voice, to producing the F0 value alone, in modal voice. The low or high F0 would be produced by relaxing or contracting the cricothyroid muscles more in syllables which had been pronounced with the constricted voice quality than in syllables that hadn't, but no longer contracting the thyroarytenoid muscles extremely at the same time. Titze (1994) shows that singers gradually shift from using the thyroarytenoid muscles to regulate F0 at the low end of their ranges to using the cricothyroid muscles at the high end. This observation suggests that if speakers who produced tense voice in syllables that originally ended in glottalic consonants stopped contracting the thyroarytenoid muscles so extremely, they would have had to compensate by contracting the cricothyroid muscles even more to keep F0 at the same high value. For these speakers then the transition from tense voice to high tone would be a matter of trying to maintain the same high F0 in modal as tense voice. The speakers who instead produced creaky voice in these syllables would simply have stopped contracting the thyroarytenoid muscles as the already relaxed cricothyroid muscles would ensure that F0 remained low.

The smaller exertion required to turn a constricted voice quality into low compared to high tone apparently supports Leer's (1999) claim that low tone is the more typical reflex, with high tone being an areal innovation in the more eastern Canadian languages. However, by hypothesis, the glottalic consonants in the ancestor of present-day high-marked languages were pronounced with stiff vocal folds. It was the original pronunciation of the consonants and not the later constriction of the vowel that made voice quality tense rather than creaky and raised F0 in the adjacent vowels. If more effort was required to produce these results, it was already being expended in the production of the consonants themselves before the contrast shifted to the preceding vowel. To shift from high F0 as an accidental byproduct of this tense voice quality to high F0 as the intended product, by shifting from combined thyroarytenoid and cricothyroid contraction to thyroarytenoid relaxation and an even higher level of cricothyroid contraction simply redistributes this effort across muscles. Recall, too, that Moore & Berke (1988) showed higher F0 values could be reached by cricothyroid than thyroarytenoid contraction; therefore, raising F0 by contracting the cricothyroid muscles may be energetically more efficient. The greater exertion therefore isn't required to produce high tone at the last stage in tonogenesis but at the first when the choice is made to pronounce the glottalic consonants stiffly rather than slackly.

3.3 *Getting to low and high tone from creaky voice alone*

In §3.2, I showed how high as well as low tone could evolve from a constricted voice quality. In this scenario, low tone evolved from the creaky variant of the constricted voice quality when the speaker contracted the thyroarytenoid muscles, but left the cricothyroid muscles relaxed and didn't stretch the covers of the vocal folds. High tone evolved from the tense variant when the speaker instead took advantage of the freedom to stretch the covers by contracting the cricothyroid muscles, too.

A high as well as a low tone could also evolve from creaky voice alone, without the speaker having to contract the cricothyroid muscles at all, at least at the stage when the contrast has shifted from the glottalic consonant to a constricted voice quality on the preceding vowel. This alternative evolutionary path depends on a characteristic acoustic property of creaky voice: following the principal intensity peak, one or two additional lower-intensity peaks or subpulses may occur within a glottal cycle in this voice quality (Figure 2; Moore & von Leden, 1958; Hollein & Wendahl, 1968; Hollein, Girard, & Coleman, 1977; Whitehead, et al., 1984; Blomgren, et al., 1998).

These subpulses apparently occur because in creaky voice the two folds' covers are so loose and their medial compression is so great that the folds can vibrate out of phase with one another (Titze, 1994). The resulting air pressure variations above the glottis are partially out of phase, and partially cancel one another when they add together. As Titze observes, this intensity modulation within a glottal cycle could affect the pitch a listener perceives. If the weaker subpulses are much less intense than the principal pulse, the listener will perceive a low pitch corresponding to the long period between the principal pulses, but if they are nearly as intense as the principal pulse, the listener will instead perceive a pitch that is two or even three times as high, corresponding to the shorter periods between subpulses. Wolfe & Ratusnik (1988) reported just such an effect: when listening the severely dysphonic voices, listeners judged pitch to be lower than in mildly dysphonic voices, most likely because secondary subpulses were much weaker relative to the principal subpulse in the severely than the mildly dysphonic voices.

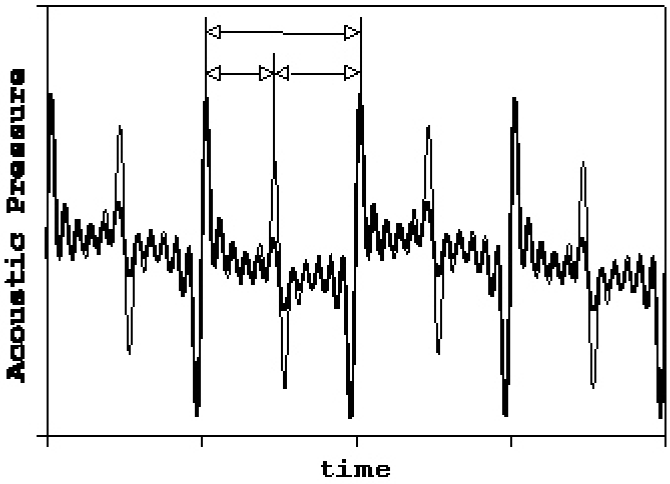


Figure 2: *A waveform where only one subpulse is intense (heavy line), whose pitch may correspond to the longer period indicated by the longer arrow, vs one with an intense second subpulse (light line), whose pitch may correspond to the shorter periods indicated by the shorter arrows.*

If speakers of one protodialect tended to produce creaky voice with secondary subpulses nearly as intense as the principal subpulse, listeners could perceive that voice quality as having a higher pitch than they would if speakers produced creaky voice with secondary subpulses markedly weaker than the principal subpulse. In this scenario, neither the glottalic consonants nor the voice quality in preceding vowels would initially have had to be pronounced so differently to produce high as well as low tonal reflexes, as required by the principal hypothesis considered in §3.2. Even so, to bring the folds' movement enough into phase to increase the secondary subpulses' intensity relative to the primary ones' and cause listeners to hear a higher pitch requires increased cover stiffness, which can only be achieved by contracting the cricothyroid. The phonetic source of high tone may therefore be the same, whether F0 is actually raised in the production of tense voice by increasing cricothyroid contraction, or only sounds like it is in creaky voice because increased cricothyroid contraction reduces the intensity differences between secondary and primary subpulses.

4. *Alternative phonetic explanations for later reversals in tone*

In §3, I've tried to explain how both low and high tones could have evolved directly from the immediate precursor of tone, vowel constriction. Constriction itself was the product of coarticulation with the laryngeal articulation of a stem-final glottalic consonant, or the ʔ coinciding with the second mora of a *Vʔ nucleus. The characteristically different F0 values of creaky and tense voice, the two ways a constricted vowel could be pronounced, were hypothesized to be the direct phonetic sources of low- and high-marked tone, respectively, in Athabaskan languages whose most recent common ancestor is the protodialect in which constriction replaced the stem-final glottalic consonants, e.g. Navajo vs Chipewyan. But, as already noted, there are also much more closely related Athabaskan languages that also differ in having low- vs high-marked tone, where a reversal of tone values must have occurred long after a glottalic articulation shifted off a stem-final consonant onto the preceding nucleus, became a constricted voice quality, and then finally tone. Furthermore, if high- and low-marked languages at first developed in geographically separate areas as Leer (1999) has argued, then high-marked languages must have recently turned into low-marked ones and vice versa.

Leer (1999) divides the Canadian and Alaskan languages that have developed tone from stem-final glottalic consonants into two groups geographically: those that originally developed low tone from this source extend north and west from central British Columbia along the west side of the Canadian Cordillera up into the drainages of the Yukon and Tanana Rivers in eastern Alaska, whereas those that instead originally developed high tone extend north and east along the east side of the Cordillera into the drainage of the Mackenzie River. This proposal entails that the high tones in Tanacross and Northern Tutchone on the Cordillera's low-marked west side are recent innovations, compared to the corresponding low tones in the other members of the Tanana and Tutchone subgroups, as is the low tone in Dogrib on the high-marked east side, compared to the corresponding high tones in the other members of the Mackenzie River subgroup.[13] Therefore, tones have reversed value recently in both directions.

A tone contrast must have replaced the earlier contrast between constricted vs non-constricted vowels much less one between nearly all glottalic vs non-glottalic stem-final consonants long before these closely related languages diverged from their recent common ancestor. However, many tonal Athabaskan languages have

---

[13] In Leer's account, Dogrib has retained the low tone that originally developed from PA glottalic consonants and it is the rest of the Mackenzie River subgroup that is innovative in replacing low with high tone.

kept a contrast between stems ending in a glottal stop and those ending in a vowel as well as one between syllables ending in a glottalic vs non-glottalic sonorant. If this glottal stop or glottalic sonorant were pronounced differently in different languages, it could reverse tone values in the syllables in which it occurred. Alternatively, listeners could have mistaken a stem's tone for a prefix's. This mistake is made likely by a combination of the crosslinguistically common tendency for tones to be realized as F0 targets late in the syllable bearing them with the particularly Athabaskan process of vowel reduction. When the prefix's tone differed from the stem's, the perceived value of the stem tone would be reversed. Neither of these mechanisms could directly reverse tone values in syllables other than those still ending in a glottalic consonant or those following a prefix with a different tone, so analogy would have to extend these limited reversals to all other syllables. These two mechanisms for reversing some tones are described in §§4.1-2, and then the analogical extensions they need to reverse tones throughout the vocabulary are described in §4.3.

4.1 *Tone reversals from persistent glottalic articulations*

The cognate sets in (3) show that in many high- and low-marked languages, syllables ending in a glottal stop still contrast with those ending in a vowel (data in 3-5 are from Krauss, 1979, this volume).

(3) *Glottalic vs modal full vowel stems *taʔ "father" vs tu· "water" in high- vs low-marked tonal languages.*

High-marked			Low-marked		
N. Tutchone	*téʔ*	*tù*	S. Tutchone	*tàʔ*	*tʃú*
Tanacross	*táʔ*	*tù·*	Upper Tanana	*tàʔ*	*tú*
Kaska	*táʔ*	*tùʔ*	Lower Tanana	*tàʔ*	*tú*
Hare	*táʔ*	*tù*	Gwich'in	*tìʔ*	*tʃų́·*
Mountain	*táʔ*	*tù*	Han	*tʃæ̀ʔ*	*tʃú·*
Bearlake	*táʔ*	*tù*	Tagish	*tàʔ*	*tú*
Slave	*táʔ*	*tù*	Tahltan	*tàʔ*	*tú·*
Beaver	*tá·ʔ*	*tʃû*	Dogrib	*tà*	*tî*
Chipewyan	*tá*	*tù*	Sekani	*tàʔ*	*tʃû*
Chilcotin	*tá*	*tù*	Sarcee	*tàʔ*	*tú*
			Navajo	*tà·ʔ*	*tó*

In many of the same tonal languages, stem-final glottalic and non-glottalic nasal sonorants also still contrast, after both reduced vowels (4) and full vowels (5).

(4) *Reduced vowel, sonorant-final stems:* *qʊnʔ *"fire" vs* *kʲən *"base".*[14]

High-marked			Low-marked		
N. Tutchone	*kʷánʔ*	*tʃàn*	S. Tutchone	*kʷə̀nʔ*	*tʃə́n*
Tanacross	*kónʔ*	*tʃàn*	Upper Tanana	*kʊ̀nʔ*	*tʃə́n*
Kaska	*kʊ́nʔ*	*tʃə̀n–éʔ*	Lower Tanana	*kʊ̀nʔ*	*tʃə́n*
Hare	*kǫ́ʔ*	*ʃı̨̀*	Gwich'in	*kǫ̀ʔ*	*tʃán*
Mountain	*kǫ́ʔ*	*tʃı̨̀*	Han	*kʷʌ̀nʔ*	*tʃʌ́n*
Bearlake	*kǫ́ʔ*	*tʃı̨̀*	Tagish	*kʷə̀nʔ*	*tʃə́n*
Slave	*kǫ́ʔ*	*tʃı̨̀*	Tahltan	*kʊ̀nʔ*	*tʃə̀n*
Beaver	*kʷə́n*	*tʃə̀n*	Dogrib	*kǫ̀ʔ*	*tʃı̨́*
Chipewyan	*kún*	*tʃı̀n*	Sekani	*kʷə̀n*	*tʃı́n*
Chilcotin	*kʷə́n*	*tʃə̀n*	Sarcee	*kų̀·ʔ*	*tʃı̨́*
			Navajo	*kǫ̀ʔ*	*tsı́n*

(5) *Full vowel, sonorant-final stems:* *kʲʷa·nʔ *"excrement" vs* *kʲa·n *"rain".*[15]

High-marked			Low-marked		
N. Tutchone	*tsá·nʔ*	*tsèn*	S. Tutchone	*są̀ʔ*	*ʃą́*
Tanacross	*tsą́·ʔ*	*tʃą̀·*	Upper Tanana	*tʼànʔ*	n.a.
Kaska	*tsǫ́ʔ*	*tʃǫ̀*	Lower Tanana	*tʼànʔ*	n.a.
Hare	*sǫ́ʔ*	*ʃǫ̀*	Gwich'in	*tʼı̀nʔ*	*tsı́n*
Mountain	*tsǫ́ʔ*	*tʃǫ̀*	Han	*tʼą̀ʔ*	*tʃą́h*
Bearlake	*tsǫ́ʔ*	*tʃǫ̀*	Tagish	*tsą́ʔ*	*tʃą́*
Slave	*tsǫ́ʔ*	*tʃǫ̀*	Tahltan	*tsàʔ*	*tʃâ·*
Beaver	*tsǫ́·ʔ*	*tʃǫ̂*	Dogrib	*tsǫ̀*	*tʃǫ́*
Chipewyan	*tsą́*	*tʃą̀*	Sekani	*tsǫ̀ʔ*	*tʃǫ̂·*
Chilcotin	n.a.	*tʃàn*	Sarcee	*tsà*	*tʃą́*
			Navajo	*tʃą̀ʔ*	*tsą́*

---

[14] The low tone in the Tahltan reflex of "base" is unexpected.

[15] The high tone in the Tagish reflex of "excrement" is unexpected, as is the falling tone in the Beaver reflex of "rain", where a low tone is expected.

If, for example, speakers of Northern Tutchone came to contract the cricothyroid as well as the thyroarytenoid muscles in pronouncing these final glottalic articulations, then those stems' nuclei would have ended with tense rather than creaky voice quality and with high rather than low F0. If speakers of Southern Tutchone, on the other hand, would have continued to pronounce the ends of these stems with only the thyroarytenoid muscles contracted, voice quality would have remained creaky, and F0 low. Tanacross speakers' pronunciations would have diverged from Upper and Lower Tanana speakers' in much the same way. Dogrib speakers would have diverged from their high-marked sisters, Hare, Bearlake, Mountain, Slave, and Chipewyan in the Mackenzie River subgroup, in the opposite direction, by ceasing to contract the cricothyroid muscles at the ends of these stems, and thereby pronouncing them with creaky voice and low F0 rather than tense voice and high F0.

The mechanism for this tone reversal is therefore identical to that for turning these languages into tone languages in the first place: stem-final glottalic articulations cause the preceding vowel to be pronounced with non-modal or 'constricted' voice quality. Because it is the same mechanism, it can produce both high- and low-marked languages, depending on whether speakers contract the cricothyroid as well as the thyroarytenoid muscles in constricting the glottis. By itself, however, this mechanism can only reverse tone values in stems ending in glottalic consonants at the time speakers opted to add or subtract cricothyroid contraction. All other syllables would be pronounced with modal voice and should be realized with the opposite tone. However, all these languages have reversed tone values in all syllables that ended etymologically in glottalic consonants, not just those which still do, including those whose final glottalic consonant was a stop or non-nasal sonorant. Furthermore, in one of the languages that has recently reversed tone values, Dogrib, glottalic consonants only still contrast with non-glottalic ones in reduced vowel stems ending in a nasal sonorant, so this mechanism wouldn't even have reversed tone values directly in Vʔ nor VVR' stems.[16] Dogrib is treated as innovative here because it's the only language in its dialect cluster that is low-marked and it's also peripheral in that cluster (cf. Leer, 2000). The way in which analogy extends this reversal to all stems ending in etymological glottalic consonants is described after the competing mechanism is discussed.

---

[16] Unless the present-day absence of a final glottalic articulation in *Vʔ and *VVR' stems is a byproduct of the tone reversal.

4.2 *Tone reversals from mistaking a stem's tone for a prefix's*

The F0 targets that phonetically realize tones typically occur very late in the syllable that is phonologically specified for that tone, often even after that syllable is over. Figure 3 shows the effects of late target realization on F0 contours in Mandarin in two-syllable utterances bearing either tone 1 or tone 3 on their first syllables and tone 1 or tone 2 on their second syllables. In these contexts, tone 1 is high level, tone 3 is low level, and tone 2 is rising. Because the first syllable's F0 targets are realized late, both tones in the second syllable begin and remain higher following tone 1 (the lines labeled 1) in the first syllable than following tone 3 (the lines labeled 3).

Late realization of tonal targets has been demonstrated both for languages in which the tones are lexical, as in the Mandarin example in Figure 3 (Xu, 1997, 1998, 1999), and for those in which they are intonational, e.g. English (Silverman & Pierrehumbert, 1990).[17] In Athabaskan itself, late realization of tonal targets has been reported by De Jong & McDonough (1993) for Navajo, where the F0 level in the second of two consecutive syllables is pulled up by a preceding high tone and down by a preceding low tone. The late realization of tonal targets has even been phonologized in Yorùbá, where a high-low sequence of tones across two syllables is realized as a high-falling sequence and a low-high sequence is realized as a low-rising sequence (Laniran, 1992); i.e. the first syllable's tonal target is realized at the beginning of the second syllable when their tones differ.[18]

---

[17] Other languages in which tonal targets are reached late in the syllable bearing them are Chichewa (Kim, 1999), Greek (Arvaniti, Ladd, & Mennen, 1998), Korean (De Jong, 1994), Mexican Spanish (Prieto, van Santen, & Hirschberg, 1995), and Oneida (Grimm, 1997).

[18] The phonologization of late target realization can also be seen in the preference of tones to shift (Clements & Ford, 1979) or spread rightward (Hyman & Schuh, 1974). In Chilcotin (Cook, 1989), a high tone spreads across the prefix domain and onto the stem itself. In Navajo (McDonough, 1999), a high but not a low tone spreads from a disjunct prefix to the immediately following conjunct prefix if its vowel is short.

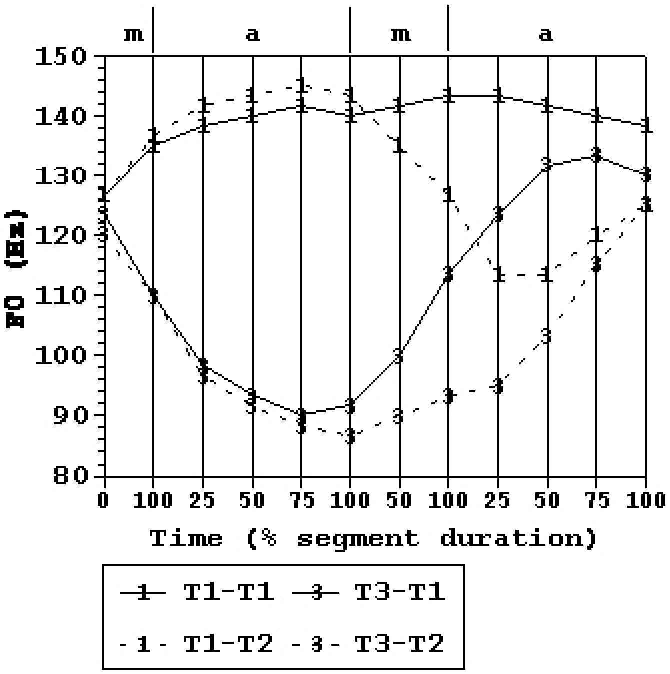


Figure 3: *F0 contours across two-syllable Mandarin utterances consisting of the segments [mama] (top) as a function of the percentage of each segment's duration for sequences of Tone 1 or Tone 3 (the numbers 1 vs 3) followed by Tone 1 or Tone 2 (solid vs dashed lines). In isolation, Tone 1 is high level 55, Tone 2 is mid rising 35, and Tone 3 is low falling rising 214. Tone 3 doesn't rise when its syllable is followed by another syllable. Redrawn from Xu (1997, Figures 3a,b, p. 67).*

Even though languages resemble one another in realizing their tones' F0 targets late, they differ in whether the realization of the F0 target is influenced by the duration of the tone-bearing syllable. In both Mandarin (Xu, 1998, 1999) and English (Silverman & Pierrehumbert, 1990; Grabe, 1998a,b), the target occurs earlier or later as the syllable is shortened or lengthened, compressing or stretching the F0 contour. In German (Grabe, 1998a,b) and most dialects of Swedish (Erikson & Alstermark, 1972; Bannert & Bredvad, 1975), on the other hand, changes in the duration of the tone-bearing syllable don't change the shape of the contour, which moves farther toward its target when the syllable is lengthened but is cut off when the syllable is shortened. These differences between languages have been attributed to whether the tonal target is coordinated with the end or the beginning of the syllable (Xu, 1998, 1999). If the F0 contour is stretched or compressed as the syllable lengthens or shortens, then the target is coordinated with the end of the syllable, but if more or less of the F0 contour is expressed, then the target is instead coordinated with the beginning of the syllable.

Because tonal targets are consistently realized late, regardless of whether they're coordinated with syllable onsets or offsets, they may never be reached in shortened syllables in languages which coordinate them with the syllable onset, because shortening cuts off the F0 contour. Under these circumstances, listeners will hear only the beginning of the F0 contour and may therefore fail to recognize the intended target (cf. Ohala, 1981). Vowel reduction would make this failure especially likely in Athabaskan languages, if it cut off the F0 contour well short of the tonal target. If vowel reduction cut off the F0 contour in the stem before its tonal target was reached, and the prefix's tonal target were reached late in the prefix syllable or even early in the stem vowel, listeners might even mistake the stem's tone for the prefix's.

This perceptual mistake will only switch the F0 value listeners attribute to the stem's tone when the prefix and stem tone differ. So, to reverse the tones permanently, prefix and stem tones must differ sufficiently often. Because the prefixes in the conjunct domain are more likely to bear unmarked than marked tone, a difference between prefix and stem tones will only be frequent if glottalic consonants weren't markedly rarer stem-finally than non-glottalic consonants in reduced vowel stems in the protolanguage.

One can estimate how often prefix and stem tones might differ, but until prefixes and stems with marked and unmarked tone are actually counted, the following estimates are merely an illustrative exercise. In the list of 270 noun stems reconstructed by Krauss (1979), the ratio of glottalic to non-glottalic stem-final consonants in reduced vowel stems is about 1:2. I have no comparable list of reconstructed verb stems, but if the noun ratio is repeated in the verbs, then stems are about half as likely to have marked as unmarked tone on a reduced vowel. If the ratio is the same for prefixes, then *ceteris paribus* in four out of nine cases prefix and stem tone will differ, while in five out of nine they'll be the same. If marked tone is disfavored even more in prefixes, the proportions of unlike tone sequences drop to 2/5, 10/27, 6/17, and 34/99 cases, i.e. hovering around a third, for marked:unmarked prefix tone ratios of 1:4, 1:8, 1:16, and 1:32, respectively. Although the stem tone is therefore increasingly likely to be the same as the prefix tone rather than different,[19] the proportions of unlike prefix-stem sequences do not become vanishingly small as marked tone becomes less frequent in prefixes. A difference between prefix and stem tone in at least one out every three cases is frequent enough that mistaking the stem's tone for the prefix's could actually have phonological consequences. So long as the mistakes are symmetric, i.e. marked stem tone is mistaken for unmarked prefix tone and vice versa, then they can reverse stem tones.

A simple shift of tones to the following syllable, as happened in Kikuyu (Clements & Ford, 1979), could not produce the reversals observed in these Athabaskan languages. Such a shift predicts incorrectly that there could be a language closely related to an existing Athabaskan language which differs only in having all its tones one syllable later, e.g. a hypothetical relative of McLeod Lake Sekani in which the low tone on the second syllable of the first prefix in _jídànííjá_ "I went in" < PA *_(ji·)daʔ-n-i·-ja·_ shifts one syllable later to produce the impossible *_jídánìjá_ in place of expected _jìdánìjà_.[20] The evolutionary path sketched here can't lead to this result because by hypothesis the perceptual mistake takes place only in stems and not in prefixes. Mistaking the current syllable's tone for the preceding syllable's is unlikely to have any phonological effect in prefixes because most of them don't bear marked tone and therefore a prefix is unlikely to differ in tone from the preceding prefix.[21]

---

[19] If stems as well as prefixes were markedly more likely to bear unmarked than marked tone, then mistaking the stem's tone for the prefix's could instead neutralize stem tone contrasts to unmarked tone rather than reversing them. This may have been how tone contrasts became vestigial or were lost altogether in some of the languages that no longer have stem-final glottalic obstruents, e.g. Koyukon, Tahltan, Carrier, and Babine.

[20] Example from Sharon Hargus, p.c. Tones are written on all syllables here for consistency with other examples, even though only low tone is phonologically active in Sekani and high tone is the default (Hargus, 1988). In this example, mistaking the stem's tone for the prefix's would have no effect because they're the same, so the predicted form is obtained by the analogical extension of tone reversal discussed in the next section.

[21] Just spreading the prefix tone, whatever it is, to the stem, won't reverse but instead only partly neutralize stem tones. In any case, typically only one prefix tone spreads synchronically in Athabaskan languages, e.g. high in Chilcotin, Navajo, and Tanacross. Furthermore, although tones usually spread to following syllables, they can also spread to preceding ones. For example, in the Hare dialect of Slave (Rice, 1990, this volume), verb stem tones all shift to the preceding syllable, neutralizing the verb stem tone to L, and in Tanacross (Holton, this volume), suffix tones dock on the stem syllable when the vowel that originally bore them is deleted. For these reasons, tone spreading from prefixes to stems probably wasn't what initiated tone reversal.

Two conditions would have to be met for tones to reverse their values by this mechanism. First, tonal targets must have been coordinated with syllable onsets rather than offsets and therefore weren't reached when the F0 contour was cut off by vowel reduction. Second, listeners misperceived the target on the stem vowel for the tone of the prefix. This mechanism, too, would not reverse all the tones that have to be reversed. The listener must analogically extend the switch of F0 values to all tones with the same phonological specification to reverse tones in: (i) reduced vowel stems preceded by a prefix with the same rather than a different tone, (ii) in full vowel stems regardless of the prefix's tone, and (iii) in unprefixed stems such as nouns.

4.3 *Extending the effects of the two mechanisms analogically*

Neither mechanism alone can reverse all the tone values that need to be reversed, the first because stem-final glottalic articulations persist only in some stems while the reversal takes place in all stems ending in etymological glottalic consonants and the second because the perceptual mistake can only occur between prefix and stem, when they differ in tone, and not elsewhere. These limited reversals must have been extended analogically.

Table 3 shows the initial reversal before persistent glottalic consonants and its analogical extension to other stems that ended in glottalic consonants in the protolanguage for a language that was originally high-marked such as Dogrib. The first row in this table shows the original high-marked pronunciation of stems ending in vowels, sonorants, and obstruents with unmarked vs marked tone, etymologically from non-glottalic vs glottalic stem-final consonants. The second row shows that a change in pronunciation of the persistent glottalic articulation reverses tone values in stems ending in vowels and sonorants but not those ending in obstruents. In the third row, this reversal is then extended analogically to stems ending in obstruents, such that any stem whose final obstruent was originally glottalic gets the low tone that appears before the persistent glottalic articulations, and any whose final obstruent was originally non-glottalic gets the high tone that now appears in stems ending in non-glottalic vowels and sonorants.

It's not that the speakers know which stems ended in glottalic and non-glottalic obstruents many generations previously. Instead, they treat all stems with the same original phonological specification for tone alike, i.e. all stems on the left side of Table 3 are originally [low] and all those on the right are originally [high]. The change in pronunciation of the persistent glottalic articulations in Vʔ and VR' stems changes the phonological value of [high] tone to [low] and is then extended analogically to all other original [high] tones, even if the phonetic conditions for the reversal aren’t met.

Stages	Unmarked Tone (Low)			Marked Tone (High)		
Original	à·	àR, ὰR	ὰK < *αK	áʔ	áR’, ά R’	άK < *αK’
Stage 1	á·	áR, άR	ὰK	àʔ	àR’, ὰR’	άK
Stage 2	á·	áR, άR	άK	àʔ	àR’, ὰR’	ὰK

Table 3: *Reversal of tone values in a high-marked language before persistent stem-final glottalic articulations (Stage 1) and its analogical extension (Stage 2) to all morphemes with etymological stem-final glottalic articulations. The changes that occur at each stage are underlined. “a” = full vowel, “α” = reduced vowel.*

Table 4 shows how mistaking a stem’s tone for a prefix’s reverses tone in sequences where prefix and stem tone are different and how this reversal is then extended analogically to all prefix-stem sequences. Mistaking the stem’s tone for the prefix’s only changes its value when its tone differs from the prefix’s. The difference in stem tone values between the first and second rows of the table in the first and third columns shows the phonological consequences of this mistake. Once [high] is reinterpreted as [low] and vice versa in stems in these contexts, those new values are extended analogically to all other morphemes with these tone values in the third row, reversing them in stems following prefixes with same tone in the second and fourth columns and in prefixes as well as stems. Again, all tones whose phonological specification is [high] must be assigned low F0 targets and vice versa. This analogical extension, like that needed by the persistent glottalic articulation mechanism, also requires listeners to sort morphemes abstractly into two distinct lexical classes by the contrast in their phonological tone specifications rather than concretely by the difference in their F0 targets. Once the F0 values attributed to some members of these two abstract classes are switched as a result of mistaking the stem’s tone for the prefix’s, F0 values are switched for all other members, completing the tone reversal. Because tones contrast to a greater extent in stems than prefixes in all the tonal languages, they are the morphemes that listeners are most likely to sort into the two abstract classes needed for the F0 switch to be extended analogically. Once extended to all stems, further analogical extension to prefixes is comparatively easy given their scanty tonal contrasts.

Stages	Low+High	High+High	High+Low	Low+Low
Original	Cà+CáC	Cá+CáC	Cá+CàC	Cà+CàC
Stage 1	Cà+CàC	Cá+CáC	Cá+CáC	Cà+CàC
Stage 2	Cá+CàC	Cà+CàC	Cà+CáC	Cá+CáC

Table 4: *Reversal of tone values in stems (CaC) following prefixes (Ca) with unlike tones (Stage 1) and analogical extension to all other morphemes with the same tone values (Stage 2). The changes that occur at each stage are again underlined.*

Whether tone reversals started with a change in the pronunciation of persistent glottalic articulations or listeners' mistaking the stem's tone for the prefix's when their tones differed, all these analogical extensions are necessary, because no Athabaskan language reverses tones only before persistent glottalic articulations or only in stems and not also prefixes, only in reduced and not also full vowel stems, only in prefixed and not unprefixed stems, or only in reduced vowel stems preceded by a prefix with the opposite tone and not also preceded by a prefix with the same tone.

At present, I don't know of any data that would permit me to choose between these two mechanisms. Both must also overcome serious difficulties: the persistent glottal stop hypothesis can't explain reversals in languages where the glottal stop hasn't persisted and the perceptual mistake hypothesis is unlikely if prefix and stem tone often don't differ. Although I can't choose between them, they're both preferable to the proposal mentioned in note 1 that tones simply reversed values, because in both hypotheses the first step to reversing tone values is phonetically motivated. Speakers may choose, emphasize, and ultimately generalize pronunciations that distinguish their speech from that of members of other speech communities, but I think they are more likely to use phonetic material already present in their speech in doing so than arbitrarily altering those pronunciations.

5. ***Concluding remarks***

In this article, I have offered three phonetic explanations, one for each step in the evolution from stem-final glottalic consonants to tone in Athabaskan. The first explained the conditions under which vowels became constricted before these consonants in the protodialect that is ancestral to the present-day tonal languages. Before the glottalic articulation of stem-final consonants was lost in this protodialect, preceding vowels were phonetically constricted as a result of coarticulation with that consonant's laryngeal articulation, so long as the vowel was reduced, or if it was full, so long as the glottalic articulation wasn't bound to the consonant, i.e. in sonorants and fricatives but not stops. In reduced vowels, even modest coarticulation was enough to change the vowel's voice quality audibly, even if the following consonant remained a stop. In full vowels, a sonorant's glottalic articulation was likely to overlap with the vowel's and constrict it, as was the glottalic articulation of a fricative derived through spirantization of a glottalic stop. However, the binding of the glottalic articulation to a stop prevented enough of a full vowel from being audibly constricted, and unmarked tone evolved instead. The second explained how both a low and high tone could have evolved directly from vowel constriction in this protodialect. A low tone evolved when vowel constriction was realized as creaky voice, whose F0 is characteristically low, and a high tone when it was instead realized as tense voice, whose F0 is characteristically high. Both voice qualities are possible realizations of vowel constriction because the thyroarytenoid and cricothyroid muscles can be contracted independently of one another, and speakers can thus control the medial compression of one vocal fold against the other by stiffening the folds' bodies independently of the longitudinal tension applied to the folds' covers. Finally, the switches of a low to a high tone or vice versa at a stage long after tone has replaced both vowel constriction and stem-final glottal consonants were explained as arising from persistent glottalic articulations in vowel- and sonorant-final stems or listeners' misperceiving the tonal target on reduced vowel stems for the tone on the immediately preceding prefix. Speakers could have recently chosen to pronounce the persistent glottalic articulations differently much as their ancestors had done when high- and low-marked languages first diverged from one another. Alternatively, listeners could have mistaken a stem's tone for a prefix's because speakers realize tonal targets late in the tone-bearing syllable and may never reach it when the F0 contour is cut off by vowel reduction. In that case, the late realization of the prefix's tonal target will be mistakenly perceived as belonging to the stem. The effects of both mechanisms for reversing tones need to be extended analogically to all tones with the same phonological value in order to reverse all tones' values.

All three of these explanations show how some speakers exercise their control of the articulators to realize a contrast in a different way than other speakers do, and thus to distinguish their own speech community from that of those other speakers (see Solnit & Kingston, 1988; Kingston & Solnit, 1989; Kingston & Diehl, 1994 for discussion). Vowel constriction distinguished the protodialect that's ancestral to the present-day tonal languages from the one that's ancestral to the present-day non-tonal languages. Realizing vowel constriction as tense vs creaky voice distinguished the speech communities within the tonal languages' ancestor from which high- vs low-marked tones evolved. This choice has remained available as a means of distinguishing speech communities up to the present day because glottalic articulations have persisted in stems ending in vowels and sonorants. For listeners to misperceive a stem's tonal target for the prefix's, speakers must have coordinated tonal targets with the stem syllable's onset rather than its offset, such that the target isn't reached when the vowel is reduced. Although the speakers' behavior is constrained by the properties of the apparatus that they use for speaking and the explanations all rely on the constraints derived from those properties, both the behavior and the explanations require that those constraints permit the phonological representation to be realized in systematically variable ways. Furthermore, to extend the recent tone reversals analogically to all other morphemes, speakers and listeners must sort tones abstractly by their phonological values and not concretely by their phonetic values or contexts.

References

Alipour-Haghighi, F. & I. R. Titze. 1985. "Viscoelastic modeling of canine vocalis muscle in relaxation". *Journal of the Acoustical Society of America* 78.1939-1943.

---. 1991. "Elastic models of vocal fold tissues". *Journal of the Acoustical Society of America* 90.1326-1331.

Alipour-Haghighi, F., I. R. Titze & A. L. Perlman. 1989. "Tetanic contraction in vocal fold muscle". *Journal of Speech and Hearing Research* 32.226-231.

Allen, E. & H. Hollein. 1973. "A laminagraphic study of pulse (vocal fry) phonation". *Folia Phoniatrica* 25.241-250.

Arvaniti, A., D. R. Ladd & I. Mennen. 1998. "Stability of tonal alignment: The case of Greek prenuclear accents". *Journal of Phonetics* 26.3-25.

Atkinson, J. E. 1978. "Correlation analysis of the physiological factors controlling fundamental frequency". *Journal of the Acoustical Society of America* 63.211-222.

Baer, T., T. Gay & S. Niimi. 1976. "Control of fundamental frequency, intensity, and register in phonation". Haskins Laboratories, *Status Report on Speech Research* SR-45/46.175-185.

Bannert, R. & A. Bredvad. 1975. "Temporal organization of Swedish tonal accents: The effect of vowel duration". Lund University, Phonetics Laboratory, *Working Papers* 10.

Bird, S. 2002. "Dakelh ejectives: Evidence for new ways of classifying sounds". 76th Annual Meeting Linguistic Society of America, San Francisco.

Blomgren, M., Y. Chen, M. L. Ng & H. R. Gilbert. 1998. "Acoustic, aerodynamic, physiologic, and perceptual properties of modal and vocal fry registers". *Journal of the Acoustical Society of America* 103.2649-2658.

Boas, F. 1911. "Tsimshian". *Handbook of American Indian Languages* ed. by F. Boas, part 1, 283-422. Washington: Government Printing Office.

Childers, D. G. & C. K. Lee. 1991. "Vocal quality factors,: Analysis, synthesis, and perception". *Journal of the Acoustical Society of America* 90.2394-2410.

Clements, G. N., & K. Ford. 1979. "Kikuyu tone shift and its synchronic consequences". *Linguistic Inquiry* 10.179-210.

Cook, E.-D. 1984. *A Sarcee Grammar*. Vancouver: University of British Columbia Press.

---. 1989. "Chilcotin tone and verb paradigms". *Athapaskan Linguistics* ed. by E.-D. Cook & K. Rice, 145-198. Berlin: Mouton de Gruyter.

Cook, E.-D., & K. Rice. 1989. "Introduction". *Athapaskan Linguistics* ed. by E.-D. Cook & K. Rice, 1-61. Berlin: Mouton de Gruyter.

Cooper, D. S., L. D. Partridge & F. Alipour-Haghighi. 1993. "Muscle energetics, vocal efficiency, and laryngeal biomechanics". *Vocal Fold Physiology* ed. by I. R. Titze, 37-92. San Diego: Singular Publishing.

Erikson, Y. & M. Alstermark. 1972. "Fundamental frequency correlates of the grave word accent in Swedish: The effect of vowel duration". Kungl Tekniska Högskolan, Speech Transmission Laboratory, *Quarterly Progress and Status Report* 2-3.

Esling, J. H., B. F. Carlson & J. G. Harris. 2002. "A laryngoscopic phonetic study of Nootka and Salish glottal stop, glottalized resonants, and pharyngeals". 76th Annual Meeting, Linguistic Society of America, and Society for the Study of Indigenous Languages of the Americas, San Francisco.

Fischer-Jørgensen, E. & H. Hirose. 1974. "A preliminary study of labial and laryngeal muscles in Danish consonant production". Haskins Laboratories, *Status Report on Speech Research*, 39/40, 231-254.

Flemming, E., P. Ladefoged, & S. Thomason. 1994. "Phonetic structures of Montana Salish". University of California, Los Angeles, *Working Papers in Phonetics* 87.1-33.

Fre Woldu, K. 1985. *The Perception and Production of Tigrinya Stops*, *Reports from Uppsala University, Linguistics Department (RUUL)*, 13.

Gobl, C. & A. Ní Chasaide. 1999. "Voice source variation in the vowel as a function of consonantal context". *Coarticulation: Theory, data and techniques* ed. by W. J. Hardcastle & N. Hewlett, 122-143. Cambridge: Cambridge University Press.

Goddard, I., 1996. "Introduction". *Handbook of Native American Indians, Language* ed. by I. Goddard, vol. 17, 1-16. Washington, DC: Smithsonian.

Golla, V. K. 1976. "Tututni (Oregon Athapaskan)". *International Journal of American Linguistics* 42.217-227.

Gordon, M. 1996. "The phonetic structures of Hupa". University of California, Los Angeles, *Working Papers in Phonetics* 93.164-187.

---. 2001. "Laryngeal timing and correspondence in Hupa". Papers in Phonology 5, University of California, Los Angeles, *Working Papers in Linguistics, v. 7,* ed. by A. Albright & T. Cho, 1-70.

Gordon, M., B. Potter, J. Dawson, W. de Reuse, & P. Ladefoged. 2001. "Phonetic structures of Western Apache". *International Journal of American Linguistics* 67. 415-448.

Grabe, E. 1998a. *Comparative Intonational Phonology: English and German*. Ph.D. dissertation, Catholic University, Nijmegen.

---. 1998b. “Pitch accent realization in English and German”. *Journal of Phonetics* 26.129-143.

Grimm, C. 1997. “Pitch accent in Oneida”. 71st Annual Meeting, Linguistic Society of America, Chicago.

Halle, M. & K. N. Stevens. 1971. “A note on laryngeal features”. Research Laboratory in Electronics, Massachusetts Institute of Technology, *Quarterly Progress Report* 101.198-213.

Han, M., & R. S. Weitzman. 1970. “Acoustic features of Korean /P,T,K/, /p,t,k/, and /pʰ,tʰ,kʰ/”. *Phonetica* 22.112-128.

Hardcastle, W. J. 1973. “Some observations on the tense-lax distinction in initial stops in Korean”. *Journal of Phonetics* 1.263-272.

Hargus, S. 1988. *The Lexical Phonology of Sekani*. New York: Garland Publishing.

Hirano, M. 1974. “Morphological structure of the vocal cord as a vibrator and its variations”. *Folia Phoniatrica* 26.89-94.

---. 1975. “Phonosurgery: Basic and clinical investigations”. *Otologia (Fukuoka)* 21.239-440.

---. 1977. “Structure and vibratory behavior of the vocal folds”. *Dynamic Aspects of Speech Production* ed. by M. Sawashima & F. S. Cooper, 13-27. Tokyo: University of Tokyo.

Hirano, M. & Y. Kakita. 1985. “Cover-body theory of vocal fold vibration”. *Speech Science* ed. by R. G. Daniloff, 1-46. San Diego: College Hill Press.

Hirano, M., S. Kurita & T. Nakashima. 1981. “The structure of the vocal folds”. *Vocal Fold Physiology* ed. by K. N. Stevens & M. Hirano, 33-41. Tokyo: University of Tokyo.

Hirano, M., K. Matsuo, Y. Kakita, H. Kawasaki & S. Kurita. 1983. “Vibratory behavior and the structure of the vocal fold”. *Vocal Fold Physiology: Current Research and Contemporary Issues* ed. by D. M. Bless, & J. H. Abbs, 26-40. San Diego: College-Hill.

Hirano, M., J. J. Ohala & W. Vennard. 1969. “The function of the laryngeal muscles in regulating fundamental frequency and intensity of phonation”. *Journal of Speech and Hearing Research* 12.616-628.

Hirano, M., W. Vennard & J. J. Ohala. 1970. “Regulation of register, pitch, and intensity of voice”. *Folia Phoniatrica* 22.1-20.

Hirose, H. & T. Gay. 1972. “Laryngeal control in vocal attack: An electromyographic study”. Haskins Laboratories, *Status Report on Speech Research* 39/40.49-60.

Hirose, H., C. Y. Lee & T. Ushijima. 1974. “Laryngeal control in Korean stop production”. *Journal of Phonetics* 2.145-152.

Hirose, H., H. S. Park, H. Yoshioka, M. Sawashima & H. Umeda. 1981. “An electromyographic study of laryngeal adjustments for Korean stops”. University of Tokyo, Research Institute for Logopedics and Phoniatrics, *Annual Bulletin* 15.31-43.

Hogan, J. T. 1976. “An analysis of the temporal features of ejective consonants”. *Phonetica* 33.275-284.

Hollein, H. 1974. “On vocal registers”. *Journal of Phonetics* 2.25-43.

Hollein, H., G. T. Girard & R. F. Coleman. 1977. "Vocal fold vibratory patterns of pulse register phonation". *Folia Phoniatrica* 29.200-205.

Hollein, H. & R. W. Wendahl. 1968. "Perceptual study of vocal fry". *Journal of the Acoustical Society of America* 43.506-509.

Holton, G. this volume. "Pitch, tone and intonation in Tanacross".

Hyman, L. M. & R. Schuh. 1974. "Universals of tone rules: Evidence from West Africa". *Linguistic Inquiry* 5.81-115.

Ingram, J., & B. Rigsby. 1987. "Glottalic stops in Gitskan: An acoustic analysis". *Proceedings of the XIth International Congress of Phonetic Sciences* 2.134-137.

De Jong, K. 1994. "Initial tones and prominence in Seoul Korean". Ohio State University, *Working Papers in Linguistics* 43.1-14.

De Jong, K., & J. McDonough. 1993. "Tone and tonogenesis in Navajo". University of California, Los Angeles, *Working Papers in Phonetics* 84.165-182.

Kagaya, R. 1974. "A fiberscopic and acoustic study of Korean stops, affricates, and fricatives". *Journal of Phonetics* 2.161-180.

Kari, J. 1990. *Ahtna Athabaskan Dictionary*, Fairbanks: University of Alaska Press.

Kim, S.-A. 1999. "Positional effect on tonal alternation in Chichewa: Phonological rule or phonetic timing?". *Proceedings of the 33rd Regional Meeting*, 245-257. Chicago: Chicago Linguistic Society.

Kingston, J. 1982. "Why tones don't come from ejectives". Ms., University of California, Berkeley.

---. 1985. *The Phonetics and Phonology of the Timing of Oral and Glottal Events*. Ph.D. dissertation, University of California, Berkeley.

---. 1990. "Articulatory binding". *Papers in Laboratory Phonology I,* ed. by J. Kingston & M.E. Beckman, 406-434. Cambridge, UK: Cambridge University Press.

Kingston, J. & R. L. Diehl. 1994. "Phonetic knowledge". *Language* 70.419-454.

Kingston, J. & D. Solnit. 1989. "The inadequacy of underspecification". *Proceedings of the 19th Meeting of the Northeast Linguistic Society,* 264-278.

Koike, Y., M. Hirano & M. Morio. 1974. "Function of the laryngeal muscles on the position and shape of the vocal cord". *Proceedings of the 16th International Congress on Logopedics and Phoniatrics*, ed. by In E. Loebell, 257-263. Basel: Karger.

Krauss, M. E. 1979, this volume. "Athabaskan tone".

Krauss, M. E., & V. K. Golla. 1981. "Northern Athapaskan languages". *Handbook of Native American Indians, Subarctic*, ed by J. Helm, vol. 6, 67-85. Washington, DC: Smithsonian.

Krauss, M. E. & J. Leer. 1981. "Athabaskan, Eyak, and Tlingit sonorants". Alaska Native Language Center, *Research Papers* 5, Fairbanks: University of Alaska.

Laniran, Y. 1992. *Intonation in Tone Languages: The Phonetic Implementation of Tone in Yorùbá*. Ph.D. dissertation, Cornell University.

Leer, J. 1979. *Proto-Athabaskan Verb Stem Variation, Part One: Phonology*. Alaska Native Language Center, *Research Papers* 1. Fairbanks: University of Alaska.

---. 1999. "Tonogenesis in Athabaskan". *Cross-linguistic Studies of Tonal Phenomena, Tonogenesis, Typology, and Related Topics* ed. by S. Kaji, vol. 1, 37-66. Tokyo:

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

---. 2001. “Shift of tone markedness in Northern Tlingit and Southern Athabaskan”. *Cross-linguistic Studies of Tonal Phenomena: Tonogenesis, Japanese Accentology, and Other Topics* ed. by S. Kaji, vol. 2, 61-86. Tokyo: Inst. Study Lgs. Cults. Asia Africa, Tokyo University of Foreign Studies.

Lindau, M. 1984. “Phonetic differences in glottalic consonants”. *Journal of Phonetics* 12.147-155.

Löfqvist, A., T. Baer, N. S. McGarr & R. Seider Story. 1989. “The cricothyroid muscle in voicing control”. *Journal of the Acoustical Society of America* 85.1314-1321.

Maddieson, I., C. L. Smith & N. Bessell. 2001. “Aspects of the phonetics of Tlingit”. *Anthropological Linguistics* 43.135-176.

Maddieson, I., R. Rajabov & A. Sonnenschein. 1996. “The main features of Tsez phonetics”. University of California, Los Angeles, *Working Papers in Phonetics* 93.94-110.

McDonough, J. M. 1999. “Tone in Navajo”. *Anthropological Linguistics* 41.503-540.

McDonough, J. M. & P. Ladefoged. 1993. “Navajo stops”. University of California, Los Angeles, *Working Papers in Phonetics* 87.151-164.

McGlone, R. 1967. “Air flow during vocal fry phonation”. *Journal of Speech and Hearing Research* 10.299-304.

McGlone, R. & T. Shipp. 1971. “Some physiologic correlates of vocal fry phonation”. *Journal of Speech and Hearing Research* 14.769-775.

Moore, D. M. & G. S. Berke. 1988. “The effect of laryngeal nerve stimulation on phonation: A glottographic study using an *in vivo* canine model”. *Journal of the Acoustical Society of America* 83.705-715.

Moore, G. P. & H. von Leden. 1958. “Dynamic variations of the vibratory pattern in the normal larynx”. *Folia Phoniatrica* 10.205-238.

Murry, T. 1971. “Subglottal pressure and air flow measures during vocal fry phonation”. *Journal of Speech and Hearing Research* 14.544-551.

Murry, T., & W. S. Brown. 1971. “Subglottal air pressure during two types of glottal activity: Vocal fry and modal phonation”. *Folia Phoniatrica* 23.440-449.

Ní Chasaide, A. & C. Gobl. 1993. “Contextual variation of the vowel voice source as a function of adjacent consonants”. *Language & Speech* 36.303-330.

Ohala, J. J. 1970. *Control and Production of Speech*, University of California, Los Angeles, *Working Papers in Phonetics*, 15.

---. 1981. “The listener as a source of sound change”. *Papers from the Parasession on Language and Behavior*, ed. by C. S. Masek, R. A. Hendrick & M. F. Miller, 178-203. Chicago: Chicago Linguistic Society.

Pike, E. V. 1986. “Tone contrasts in Central Carrier (Athapaskan)”. *International Journal of American Linguistics* 52.411-418.

Pinkerton, S. 1986. “Quichean (Mayan) glottalized and nonglottalized stops: A phonetic study with implications for phonological universals”. *Experimental Phonology* ed. by J. J. Ohala & J. J. Jaeger, 125-139. Orlando: Academic Press.

Prieto, P., J. van Santen & J. Hirschberg. 1995. “Tonal alignment patterns in Spanish”. *Journal of Phonetics* 23.429-451.

Rice, K. 1990. “Predicting rule domains in the phrasal phonology”. *The phonology/syntax connection* ed. by S. Inkelas & D. Zec, 289-312. Chicago: University of Chicago Press.

---. 1993. “A reexamination of the feature [sonorant]: The status of ‘sonorant obstruents’”. *Language* 69.308-344.

---. 1995. “The representation of the perfective suffix in the Athapaskan language family”. *International Journal of American Linguistics* 61.1-37.

---. 1997. “A reexamination of Proto-Athabaskan *y*”. *Anthropological Linguistics* 39.423-436.

---. this volume. “Prominence and the verb stem in Slave (Hare)”.

Rice, K. & S. Hargus. this volume. “Introduction”.

Rubin, H. J. 1963. “Experimental studies on vocal pitch and intensity in phonation”. *Laryngoscope* 73.973-1015.

Sands, B., I. Maddieson & P. Ladefoged. 1993. “The phonetic structures of Hadza”. University of California, Los Angeles, *Working Papers in Phonetics* 84.67-87.

Sapir, E. 1922. “The Takelma language of Southwestern Oregon”. *Handbook of American Indian Languages* ed. by F. Boas, part 2, 1-296. Washington, DC: Government Printing Office.

---. 1938. “Glottalized continuants in Navajo, Nootka, and Kwakiutl (with a note on Indo-European)”. *Language* 14.248-274.

Shipley, W. F. 1978. “Native languages of California”. *Handbook of Native American Indians, California* ed. by R. F. Heizer, vol. 8, 80-90. Washington, DC: Smithsonian.

Shipp, T. & R. E. McGlone. 1971. “Laryngeal dynamics associated with voice frequency change”. *Journal of Speech and Hearing Research* 14.761-768.

Silverman, K. E. A. & J. B. Pierrehumbert. 1990. “The timing of prenuclear high accents in English”. *Papers in Laboratory Phonology I* ed. by J. Kingston & M. E. Beckman, 72-106. Cambridge, UK: Cambridge University Press.

Solnit, D. & J. Kingston. 1988. “Voiced-high splits: Putting the blame on sonorants”. 63rd Annual Meeting of the Linguistic Society of America, New Orleans.

Steriade, D. 1997. “Phonetics in phonology: The case of laryngeal neutralization”. Ms., University of California, Los Angeles.

Stevens, K. N. 1977. “Physics of laryngeal behavior and larynx modes”. *Phonetica* 34.264-279.

Story, B. H. & I. R. Titze. 1995. “Voice simulation with a body-cover model of the vocal folds”. *Journal of the Acoustical Society of America* 97.1249-1260.

Story, G. 1989. “A report on the nature of Carrier pitch phenomena: With special reference to the verb prefix tonomechanics”. *Athapaskan Linguistics* ed. by K. Rice & E.-D. Cook, 99-144. Berlin: Mouton de Gruyter.

Tarpent, M.-L. 1987. *A Grammar of the Nisgha language*, Ph.D. dissertation, University of Victoria.

Thrainsson, H. 1978. “On the phonology of Icelandic preaspiration”. *Nordic Journal of Linguistics* 1.3-54.

Titze, I. R. 1994. *Principles of Voice Production*. Englewood Cliffs, NJ: Prentice Hall.

Titze, I. R., E. Luschei & M. Hirano. 1989. “The role of the thyroarytenoid muscle in regulation of fundamental frequency”. *Journal of Voice* 3.213-224.

Tuttle, S. G. 1997. “Acoustic realizations of laryngeal contrasts in Minto Athabaskan obstruents”. Athabaskan Languages Conference, Eugene, OR.

---. 1998. “Ejectives in Tanana Athabaskan”. *Journal of the Acoustical Society of America* 103.3086. (Abstract)

Whitehead, R. W., D. E. Metz & B. H. Whitehead. 1984. “Vibratory patterns of the vocal folds during pulse register phonation”. *Journal of the Acoustical Society of America* 75.1293-1297.

Wolfe, V. & D. Ratusnik. 1988. “Acoustic and perceptual measurements of roughness influencing the judgments of pitch”. *Journal of Speech and Hearing Disorders* 53.15-22.

Wright, R., S. Hargus & K. Davis. 2002. “On the categorization of ejectives: Data from Witsuwit’en”. *Journal of the International Phonetic Association* 32. 43-77.

Xu, Y. 1997. “Contextual tonal variations in Mandarin”. *Journal of Phonetics* 25.61-83.

---. 1998. “Consistency of tone-syllable alignment across different syllable structures and speaking rates”. *Phonetica* 55.179-203.

---. 1999. “Effects of tone and focus on the formation and alignment of f0 contours”. *Journal of Phonetics* 27.55-105.

Young, R. W. 1983. “Apachean languages”. *Handbook of Native American Indians, Southwest* ed. by A. Ortiz, vol. 10, 393-400. Washington, DC: Smithsonian.

Zemlin, W. R. 1988. *Speech and Hearing Science*. 3rd ed. Englewood Cliffs, NJ: Prentice Hall.