10-2014

Representational Implications of the Phonologization of Contour Tones

Benjamin Kirkland Macaulay
Graduate Center, City University of New York

REPRESENTATIONAL IMPLICATIONS

OF THE PHONOLOGIZATION OF CONTOUR TONES

by

BENJAMIN K. MACAULAY

A master's thesis submitted to the Graduate Faculty in Linguistics in partial fulfillment of the requirements for the degree of Master of Arts, The City University of New York

2014

This manuscript has been read and accepted for the
Graduate Faculty in Linguistics in satisfaction of the
Dissertation requirement for the degree of Master of Arts.

Juliette Blevins

Date

Thesis Advisor

Gita Martohardjono

Date

Executive Officer

THE CITY UNIVERSITY OF NEW YORK

Abstract

REPRESENTATIONAL IMPLICATIONS
OF THE PHONOLOGIZATION OF CONTOUR TONES

by

Benjamin K. Macaulay

Advisor: Professor Juliette Blevins

This thesis bridges accounts of tonogenesis to representations of contour tones in the dominant framework, autosegmental phonology. Accounts of tonogenesis reference phonetic features and structures that are unable to be represented in autosegmental phonology. As these features are required for the phonologization of contour tones, it is argued that they must also receive some representation in the synchrony. This is done under the Evolutionary Phonology framework of sound change. An attempt to reconcile the disparity between perceived phonetic features and synchronic structure is made by discussing the implications of the sequencing of targets within autosegmental contours. This innovation would better reflect the tonal systems of the world's languages as well as other asymmetries in segmental phonology.

Acknowledgements

I have many people to thank for the guidance and inspiration it took to complete this thesis. Working under Juliette Blevins has been a great experience. In a field unwilling to question its own theory it is a breath of fresh air having an adviser willing to engage the material honestly. Chilin Shih made a great effort to correspond with me about her work in tone and speaker intuitions. Pattie Epps selflessly provided me with invaluable data.

Secondly, I must thank Lucien Szpiro, as without his employment I would have never made it through my graduate studies. Having a flexible work situation (and access to a photocopier) made a great difference in my ability to deal with my workload.

My peers at the Graduate Center have also been an invaluable resource. Without Toni Foster I would have no idea how to format this thesis. I am thankful to Lauren Spradlin for assuring me that my work has not driven me mad (or alternatively, that it has driven us mad together).

Finally, I would like to thank Nigel Massey for steering me towards linguistics when I didn't realize it was "a thing." I am ecstatic to finally find an institution where my obsessions and neuroticism are productive. But to fully list the ways he has inspired me would be outside the scope of these Acknowledgements.

Table of Contents

List of Tables

List of Figures

1 – Introduction

Theories of phonology and distinctive features have long focused on the synchrony. In these theories, features and types of representation are omitted from these theories of phonology if they can be accounted for by some other means (in the name of efficiency). However, this tendency is rooted in the assumption that all features and representations necessary for an accurate theory of phonology will be referenced by synchronic phonological rules and alternations. This paper examines an additional source of features necessary to explain the sound patterns of the world's languages: sound change.

This paper will focus on tone, more specifically contour tones. Contour tones have posed a problem for linguists working with synchronic grammars: the languages with the richest tonal inventories (such as those in Asia and Central America) often lack the kinds of morphological and phonological alternations that linguists have traditionally used to determine the features and representations necessary to account for patterns in other kinds of speech sounds. Because of this, features and representations that may be active in these languages are overlooked for autosegmental workarounds that are often ad hoc.

However, tone is also special in that a language's use of contrastive tone can often be traced back to its advent. This involves tonogenesis, the process by which languages without lexical tone develop lexical tone. This paper seeks to explore accounts of tonogenesis, in order to investigate whether the representation of contour tones in the dominant autosegmental framework can account for the structure and phonetic features referenced during these tones' phonologization.

The way this will be done is by adopting the Evolutionary Phonology framework of sound change (Blevins 2004). By expanding this framework to accommodate supralaryngeal features and autosegmental contours, the adequacy of current representations of contour tones and the limited distinctive features said to govern them will be examined.

Section 2 will give an overview of tonal phonetics as relevant to tonogenesis and the discussed case studies later in the paper. Section 3 approaches tone from a phonological perspective and outlines the peculiarities in current frameworks of tonal phonology as they relate to phonetic accounts of tonogenesis. Section 4 takes the reader through some of the different processes which result in phonological contour tones and Section 5 concludes by addressing the relationship between this historical data and the representation of contour tones in autosegmental phonology with a formal discussion of tonogenesis under Evolutionary Phonology.

2 – Phonetics of laryngeal features and tonogenesis

Tonogenesis mainly involves the phonetics of laryngeal features, not only their perceptual and acoustic properties but in certain cases their articulation as well. This is not to say that supralaryngeal features and articulations don't come into play; anything that affects pitch or positioning in the larynx can be relevant to tonogenesis. (Examples of this will be provided in 2.5.)

This section will provide a background in the laryngeal phonetics relevant to tonogenesis and how tonogenesis occurs from a purely phonetic perspective. 2.1 will outline the anatomy of the larynx, 2.2 will discuss the manipulation of tone, 2.3 will explain the effects of phonation in the absence of supralaryngeal gestures and 2.4 will show the effects of phonation when combined with other articulations. 2.5 will discuss tonogenesis from a purely phonetic stance, and its relation to laryngeal and supralaryngeal features.

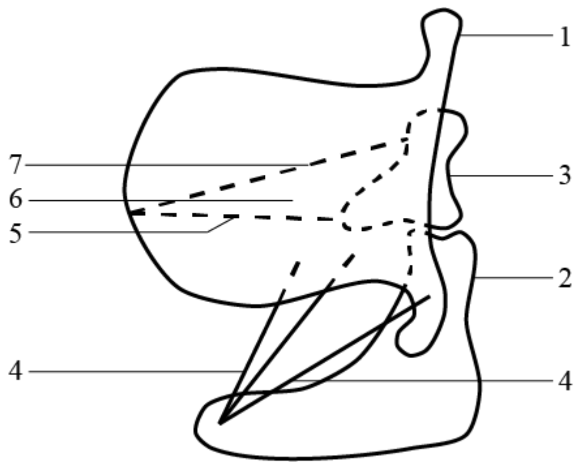
2.1 – Anatomy of the larynx

The larynx contains four cartilages: the ring-shaped cricoid cartilage, the large crown-shaped thyroid cartilage that sits above it and the two small wing-shaped arytenoid cartilages that sit above the rear of the cricoid cartilage where there is a break in the thyroid (Laver 1980:99).

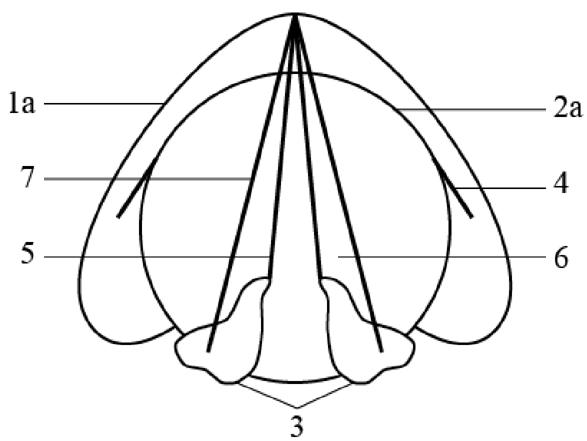
Between the arytenoid cartilages and the front of the thyroid are the vocal folds, as well as the ventricular (or "false") vocal folds, which sit above them. The glottis is the space that opens between the vocal folds (sometimes including the vocal folds themselves); the space between the arytenoid cartilages is the cartilaginous glottis while the space between the vocal folds from the thyroid up until the arytenoids is the ligamental glottis (Laver 1980:105).

The anatomy of the larynx can be seen in Figure 1, from Laver 1980:102:

Figure 1: structure of the larynx (Laver 1980:102)



side view



overhead view

Legend:

1. thyroid cartilage

1a. edge of thyroid

2. cricoid cartilage

2a. edge of cricoid

3. arytenoid cartileges

4. cricothyroid muscle

5. edge of glottis

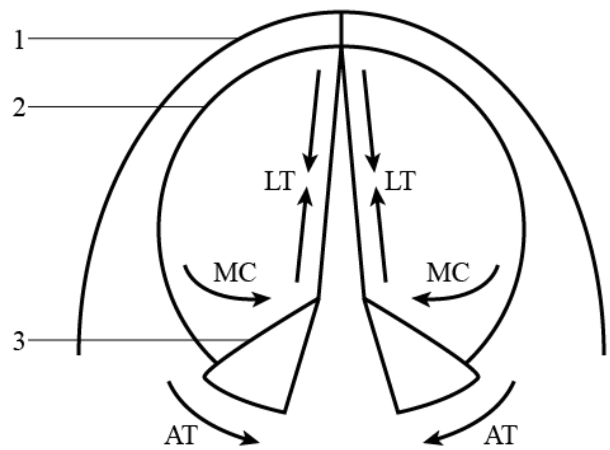
6. Ventricle of Morgagni

7. ventricular vocal folds

Various muscles connect the cartilages, and laryngeal settings in which different combinations of these muscles are contracted will have effects on the speaker's voice quality called phonation. Laver 1980:108-9 reduces the gamut of these settings to three variables: longitudinal tension, medial compression and adductive tension.

Longitudinal tension is a measure of tension of the cricothyroid muscle, which tips the front of the cricoid up, tilting the end connected to the arytenoid cartilages back and stretching the ligamental glottis. Adductive tension is the width of the cartilaginous glottis, controlled by the interarytenoid muscles which connect the arytenoid cartilages. Medial compression is the inward tension of the vocal folds, closing the ligamental glottis. This is achieved by tension of the interarytenoid muscles (overlapping with adductive tension in this regard) as well as other muscles that manipulate the arytenoids such as the lateral cricoarytenoid muscles and the thyroarytenoid muscles which make up the vocal fold body.

Figure 2: laryngeal parameters relevant to phonation (Laver 1980:109)



Legend:

AT adductive tension

MC medial compression

LT longitudinal tension

1. thyroid cartilage

2. cricoid cartilage

3. arytenoid cartilages

In addition to the cartilages and muscles that compose the larynx itself, there are tissues surrounding the larynx that aid in its control. Above the larynx is the hyoid bone, and the suprahyoid muscles which connect this bone to the chin will raise the larynx when contracted. Similarly, the strap muscles which extend towards the collarbone will lower the larynx.

The “default” laryngeal setting for speech is modal voice, which exhibits the full pitch range of speech (Laver 1980:110). Modal voice occurs under moderate adductive longitudinal and adductive tensions and moderate medial compression.

2.2 – Tone

Whenever there is periodic vibration of the vocal folds (phonation), speech will have a fundamental frequency (F_0), also known as "pitch" or "tone." "Pitch" is a perceptual term, describing the acoustic property of speech perceived by speakers that corresponds to F_0.[1] "Tone" refers to the manipulation of the F_0 of speech, either to convey meaning at the phrasal level (for example, final rises that signal questions in English and many languages cross-linguistically), or to distinguish between lexical items. Languages that do the latter have 'lexical tone,' and these lexically-specified tones are overlaid on the larger phrasal intonation (Fujisaki et al. 2007:236).

Speakers can raise their fundamental frequency in a number of ways. The cricothyroid muscle can manipulate the thyroid in such a way that the vocal cords lengthen and tense, either by translation of the thyroid, pushing it forward by contracting the *pars oblique* of the cricothyroid, or by rotation, pushing tipping the front of the thyroid downward by contracting the cricothyroid's *pars recta* (Fujisaki et al. 2007:233). Contraction of the suprahyoid muscle to raise the larynx will also raise F_0 (Stevens 2000:251).

F_0 can also be actively lowered by speakers. This is achieved by stabilizing the hyoid bone with the sternohyoid muscle, which allows for the thyrohyoid muscle to rotate the thyroid in the opposite direction as the cricothyroid. This reduces the length of the vocal cords and their tension, lowering frequency (Fujisaki et al. 2007:234-6). The larynx itself can also be lowered with the strap muscles to lower frequency (Stevens 2000:251).

[1] While pitch is properly a perceptual term, it is also often used to refer to the F_0 of speech in general, and in terms such as "pitch accent" which are not strictly perceptual. Likewise, in this paper "pitch" does not necessarily refer only to the perceptual property of F_0 but also in a general acoustic sense.

In order to denote what part of the pitch range a tone is in (abstracting away from coarticulation with prosodic melodies and only considering the range of pitch used for speech), I will use two main notations:[2] H/M/L to denote high/middle/low tones, and a 1-5 scale where 1 is the lowest pitch and 5 the highest, following Chao 1930. Contour tones where a tone rises, falls or exhibits a combination of the two can be denoted by a sequence such as HL for falling or 35 for high rising.[3]

There is a further property of tones that must be defined for the purposes of this analysis. To compare for example a low-falling and high-falling tone (ML and HM or 31 and 53), these tones share neither a common starting point nor endpoint in the pitch range. However, there is a clear commonality between them, namely that each tone falls in pitch over the course of its contour. This property, the vertical movement of the tone within the pitch range over time, I will refer to as the *tone differential*, after the similar mathematical function that isolates the slope of a function without reference to its height in the space. The tone differential stands in contrast to the specific height in the pitch range (after abstracting away from overarching prosodic tones etc.), this I will simply refer to as *tone height*. Level tones are tones with a null tone differential, contour tones have a non-null tone differential (rising, concave etc.). Contour tones may still be distinguished by tone height, for example languages like Cantonese have an upper and lower tone register, such that high-rising MH and low-rising LM are distinguished by tone height but share a tone differential. Tone registers will be discussed further in 3.3.

---

[2] In addition to the tone symbols and diacritics used by the International Phonetic Alphabet.

[3] Some authors use R and F to denote rising/falling in letter notation instead of HL/LH. This is useful for analyses of contour tones as unitary, as opposed to separable level tones (as will be discussed in 3.3 and 3.4). However, the use of HL/LH in this paper does not assume the validity of the latter analysis, and R/F will only be used when a unitary contour tone is necessary.

For the sake of the reader I will recap the terminology used for this paper:

Table 1: Properties of tone

tone height	The relative F_0 within the range of F_0 relevant to speech. Transcribed either by Chao numerals (1-5 where 1 is the bottom and 5 is the top of the pitch range), letters such as H for high tone and L for low tone or equivalent IPA diacritics.
tone register	The status of tone as either being in the upper or lower half of the F_0 range relevant to speech. The property that contrasts a high-rising with a low-rising tone is the register. Register is used in some feature analyses.
tone differential	The direction of change over time in the F_0 contour, equivalent to the "slope" of an equation on a graph. Rises and falls in pitch have positive and negative tone differential values, respectively. Level tones have a zero value for the tone differential. The feature shared by a high-rising and a low-rising tone is their positive tone differential.

2.3 – Phonation

Through manipulation of the larynx as sketched in 2.1, variations of phonation other than modal voice can be achieved. These phonation types are not only often bundled with specific tones in "register complex" languages (or the main cues in non-tonal register languages), but play a pivotal role in tonogenesis.

2.3.1 – Breathy voice

Breathy voice is a laryngeal setting that involves little adductive tension, weak medial compression and low longitudinal tension (Laver 1980:133). The result is a glottis spread wider than that of modal voice, but such that the vocal cords are still able to vibrate periodically, albeit with less frequency. In Distinctive Feature Theory it is denoted by [SPREAD GLOTTIS] and [+VOICE].

This less frequent vibration gives breathy voice a low F_0. Because of this, many languages have reinterpreted breathy voice as low tone. An example of this is Punjabi, where breathy voiced consonants have merged with unaspirated voiceless stops, leaving only tonal reflexes behind (L tone associated with the following two syllables, or H tone for the two previous if the consonant was stem-final[4]) (Gill & Gleason 1969, Ohala 1973:11).

2.3.2 – "Laryngealization"

The terms "laryngealization" and "glottalization" are often used in the literature to describe phonation with some sort of constriction in the glottis (thus given the designation [CONSTRICTED GLOTTIS] in Distinctive Feature Theory). However, "laryngealization" does not refer to one single laryngeal setting, but at least two: creaky voice and harsh voice, the latter described alternatively as stiff voice (Ladefoged & Maddieson 1996:55-7) or tense voice.

Creaky voice involves strong medial compression and adductive tension (Hollien et al. 1966:247), often also a raising of the larynx (Thurgood 2008:15, Matisoff 1973:76); this can result in the vocal folds touching the ventricular folds, possibly creating one vibrating mass (Hollien et al. 1966:247). Ladefoged 1996:53 notes that the cartilaginous and ligamental glottides vibrate separately during creaky voice. If these vibrations are out of phase, they can create the illusion of a doubling of glottal pulsation (Ladefoged & Maddieson 1996:54 Fig3.3, demonstrated in Fula (Niger-Congo) during voiced creaky/modal stops; Kingston 2005:164-5, Titze 1994). Otherwise, creaky voice has a lowering effect on F_0 (Laver 1980:122).

[4] The H tone is an unexpected reflex of the low-frequency breathy voiced consonants. One possible explanation could be that these syllables had a higher relative pitch than the coda $*D^h$, and the negative tone differential was phonologized as an H tone on preceding syllables. An in-depth discussion of Punjabi is out of the scope of this paper; phonologization of the tone differential will be discussed later in the paper in the context of contour tones.

Harsh voice has a similar character as creaky voice, however while creaky voice is generally below 100Hz, harsh voice is above 100Hz (Ladefoged 1996:56, Laver 1980:130). It is achieved through higher medial compression and adductive tension (Laver 1980:131).

Creaky voice and tense voice share properties other than F_0: they both are articulated with heightened medial compression and adductive tension, and they both commonly originate from or are allophones of glottal stops (discussion in 2.4.2 and Garallek 2013). However, as the effects of laryngealization on F_0 are relevant to this paper, creaky and tense voice must be distinguished. Languages will associate low tones to environments with the low-frequency creaky voice, and high tones to environments with the high-frequency tense voice. Kingston 2005:163-4 mentions the creaky/tense distinction as a possible source of opposing tone values in Athabaskan word classes stemming from some laryngealized source.[5] Creaky voice is found on two of the low register tones in Vietnamese (the third having optional breathy voice), and the Danish *stød* (realized as creaky voice or glottal stop) is accompanied by a low tone (Gussenhoven 2004). Another example of tense voice is in the register language Takhian Thong Chong (Pearic, Mon-Khmer), where the tense voice register has a much higher F_0 than the other three registers (modal, breathy and a breathy-tense contour) (DiCanio 2009:14).

2.4 – Non-phonatory laryngeal features

Laryngeal settings that do not produce periodic vocal cord vibration ("nil phonation") can also have effects on pitch in their phonetic environment. These include the glottal segments [h] and glottal stop, as well as laryngeal features on obstruents and consonants produced with glottal airstream mechanisms.

[5] The tone reversal in Athabaskan will be discussed in-depth in 4.2.3.

2.4.1 – [h]

[h] is the result of highly turbulent airflow through a widely abducted glottis (Laver 1994:189). Turbulence of airflow begins at 2-300 cc/s, measurements of [h] are around 1000 cc/s (Catford 1997:95). The status of [h] in featural terms is ambiguous, however. Because of the turbulence in the glottis, [h] is often referred to as a glottal fricative, as it is on the IPA chart. However, as [h] lacks a constriction in the oral cavity, the tongue's position can leave a trace of the vowel quality that would result under phonation. From this perspective, [h] is identical to a voiceless vowel. (ibid.) Lieberman 1997:180 notes that [h] on a spectrogram resembles a "noise-excited version of the vowel [a]," presumably as his sample was taken from a speaker whose tongue was resting in the [a] position. He concludes that [h] can be regarded as "vowels [with place features from surrounding vowels] excited by noise excitation generated at, or near the level of the glottis." These traces of vowel quality must also be at least somewhat perceptible to speakers: Japanese has minimal pairs where voiceless allophones of high vowels are contrasted for backness (Ladefoged & Maddieson 1996:315).

While [h] itself does not exhibit periodic vocal cord vibration, it can have an effect on tone. Thakhian Thong Chong's registers use phonation type as the primary cue and thus do not contrast for tone, however syllables in the modal/breathy registers have a rising tone if [h] is in the coda (DiCanio 2009:5-6). Punjabi's high tone on stems ending in *-h* is possibly also an example of this (Thurgood 2008:14, Ohala 1973:11).[6] Thurgood 2008:14-15 also notes languages where a final [h] has resulted in a low tone, positing that these processes involved a stage of breathy voice, as opposed to the "abrupt" or "non-breathy" voiceless [h]. The motivation

[6] Unless this **-h* was realized as the voiced [ɦ] at the time of tonogenesis, patterning with the breathy voiced consonants (see Footnote 4).

for high tones associated with [h] may stem from their high rate of airflow, which is correlated with higher F_0 when controlling for other factors (Titze 1991:136, Hombert 1976:214).

2.4.2 – Glottal stop

A glottal stop is produced by fully adducting the vocal folds, continuously exerting pulmonic egressive effort culminating in a surge of transglottal airflow as the glottis reopens (Laver 1994:187-8). In practice, speakers do not always fully produce a stop when a glottal stop is present phonologically: the realization of a glottal stop can be anywhere from a full stop to laryngealized phonation on a vowel (Garellek 2013:1-2, Gerfen & Baker 2005:332, Pierrehumbert & Frisch 1997:9-10).

Experimental evidence has shown that glottal stops are perceived as a dip in pitch (Garellek 2013:55-7). Hillenbrand & Houde 1996's perception experiment manipulated F0 in English utterances like "oh-oh" [ˈoʔo] and found that the strongest cue for perception of glottal stops for English speakers was a drop in F_0. Pierrehumbert & Frisch 1997 ran a production experiment with English speakers in pairs like "heavy yoke" vs. "heavy oak" ([ˌhɛviˈ(j)oʊk] and [ˌhɛviˈʔoʊk] respectively). They also found a dip in F_0 as long as .14 seconds, dropping to an average of 27Hz for male speakers and 43Hz for female speakers.

2.4.3 – Voice Onset Time

Voice Onset Time (VOT) is a continuum of the amount of time between release of an oral obstruent and the onset of phonation. (Truly) voiced consonants have a negative voice onset time (or at least voicing through the duration of the oral closure). Aspirated consonants have a large positive VOT, with a notable delay between the closure interval and start of vocal cord vibration maintained by abduction of the glottis (Ladefoged & Maddieson 1996:95). Voiceless unaspirated consonants do not have a large positive VOT nor phonation during the full closure interval.

Rothenburg 2009:6-8 notes that this does not exhaust the permutations of consonant articulation and glottal abduction, for example preaspirates, consonants specified for breathy voice which abduct the glottis without a voice onset, or partial voicing during the closure interval.

VOT values can affect pitch in various ways. Voiced consonants have a low F0 as a lowering of the larynx and expansion of the vocal tract is required to maintain voicing during a closure interval (Ladefoged & Maddieson 1996:95, Hombert et al. 1979:40-1, Rothenburg 2009:6). Voiceless unaspirates have also been reported to slightly raise F_0 (Hyman & Schuh 1974:110).

Aspiration does not have one uniform effect on fundamental frequency, however. Hombert 1975b reported a higher F_0 at voice onset in Korean syllables, presumably from the higher airflow (see 2.4.1). This would predict aspirates generating high tones, which is borne out in the languages of mainland New Caledonia (Rivierre 1993). However, Hombert 1976 shows evidence from English and French that an exact correlation between specific VOT values and effects on pitch is unlikely. French and English both have voiceless unaspirates, however the French unaspirates behave more like English aspirates than English unaspirates. Xu & Xu 2003 note that Hombert 1976's French and English data may not be comparable and show evidence from Mandarin of vowels having a *lower* F_0 after aspirated consonants than unaspirated ones. They posit two possible reasons for aspirates to lower pitch: 1) subglottal pressure (Ps) will be *increasing* at the end of an unaspirated closure interval but *decreasing* after the high-airflow aspirates, or 2) the voiceless unaspirates are articulated with stiffness in the vocal folds causing a higher pitch in the following vowel (Xu & Xu 2003:3-4). Of note also is that Rothenburg 2009:5-6 shows a short period of breathy voice in the transition from nil to modal phonation at voice onset; he mentions that this might be perceived by speakers in English but does not substantiate the claim.

2.4.4 – Glottal airstream mechanisms

Adduction of the glottis is also used in non-pulmonic airstream mechanisms. In ejective consonants, the glottis is adducted and the larynx raised, causing the air pressure in the vocal tract to double and release into a high-amplitude stop burst (Ladefoged & Maddieson 1996:77-8). For implosive consonants, the glottis is adducted but lowered instead, lowering the air pressure before the stop burst. A key difference is that while ejectives must be voiceless during the closure interval,[7] the downward laryngeal motion of implosives causes some spontaneous vibration of the vocal folds.

Greenberg 1970:132 reports that both ejectives and implosives fail to lower pitch. He notes the status of implosive [ɓ] in Bassa (Kru, Liberia), where it is part of a phonological class that raises certain tones. Hyman & Schuh 1974:110 also reports a high rise in F_0 after implosives. DiCanio 2012:163 describes the effects of ejectives on pitch in Itunyoso Trique: carefully-pronounced ejectives raise F_0 but "slack ejectives" result in creaky voice and a low F_0 on the following vowel.

2.5 – Tonogenesis

Tonogenesis is the process by which languages develop lexical tone. This happens when a contrast in some feature that affects F_0 (such as the laryngeal features discussed in 2.3 and 2.4) is reanalyzed as a contrast in the tones that govern those F_0 values. Articulation also plays a role, for example breathy voice and [h] share an articulation that abducts the glottis but have opposing

[7] Ladefoged & Maddieson 1996:80 reports contrastively voiced ejectives in two languages, but the ejectives are prevoiced. Raising the larynx as required for ejectives does not seem possible during phonation.

effects on pitch, but a realization of /h/ as breathy voice (or such a sound change) is not unheard of; the falling tones in Vietnamese from *-*h* may have had an intermediate breathy stage where the abducted glottis and phonation overlapped (Thurgood 2008:14-15). The larynx-raising gesture that accompanies creaky voice (Matisoff 1973:76) could also raise F_0 if not accompanied by the other articulations that govern creaky voice, while creaky voice itself lowers F_0.

Interestingly, not all phonetic features that affect F_0 seem to act as the catalyst for tonogenesis. Vowel height has a clear positive correlation with F_0 (Hyman & Schuh 1974, Stevens 2000:288, Laver 1994:455), however languages do not seem to develop contrasting tones based on an (earlier) distinction in vowel height (Hombert 1978b:96-7). There is evidence of a new extra-high tone developing in Shinasha (Omotic, Ethiopia) on high vowels /i ɨ u/ in syllables already specified for high tone (Tesfaye & Wedekind 1994:13-14). However, this language already had lexical tone; accounts of *de novo* tonogenesis seem to involve laryngeal features. In fact, it is more common for the reverse to happen, that is, languages with register distinctions splitting or redistributing vowel systems. Thurgood 1999:202 describes this in Haroi (Chamic), where a three-way register system including tense and breathy voice lowered and raised vowels respectively, causing an increase in the vowel inventory.

Not all present laryngeal features will participate in a tonogenesis process either; while some languages have developed a tonal split on the basis of aspiration (see 2.4.3), Chinese, Thai and Vietnamese have retained a contrast in onset aspiration without affecting the evolution of contrastive tones. Lhasa Tibetan loanword phonology from Mandarin completely ignores the original Mandarin tones and laryngeal features; assigning loanwords H if their onsets are [-SON] or L for a [+SON] onset (Hsieh & Kenstowicz 2008:282). In reality all that can be said about these laryngeal, manner of articulation and airstream features is that they are more or less likely to generate a split in the tonal inventory, and are more or less likely to affect tone in a certain way. Hyman & Schuh 1974:110 ranks these properties of consonant series as the following (examples mine from discussions in 2.3-2.5, table taken from discussion in Greenberg 1970:132-3):

Table 2: Ranking of tonal coarticulations by consonant series

		Examples:
More likely to raise tone	implosives	raising, Bassa (Greenberg 1970)
	aspirated	raising, Mainland New Caledonia (Rivierre 1993)
		lowering, Mandarin (Xu & Xu 2003)
	voiceless unaspirated	raising, Mandarin (Xu & Xu 2003)
		lowering, Mainland New Caledonia (Rivierre 1993)
	sonorants	lowering, Lhasa Tibetan (Hsieh & Kenstowicz 2008)
	voiced unaspirated	lowering, Bassa (Greenberg 1970)
More likely to lower tone	breathy voiced	lowering, Punjabi (Ohala 1978)

Another peculiarity of tonogenesis is its relationship with phonological domains. Haudricourt 1954 famously proposed that level tones arise from conditions in the syllable onset and contour tones from conditions in the coda. While this claim has since been rejected (Thurgood 2008, Hombert 1984), experimental data shows that contour tones are better perceived towards the end of the syllable, especially for falling tones due to a "masking effect" where the phonetic information in the onset obscures tone perception (Hombert 1975a:226, Silverman 1995:9) as well as the natural delay in F0 that arises from the timing of laryngeal vs. supralaryngeal articulation (Xu 1999:1883). Not surprisingly then, while both level and contour tones have been shown to arise from the rime, cases of a contour tone directly attributed to an onset feature are rare. This is shown in the following table for reference:[8]

[8] A grain of salt must be taken when considering the "environments" as laid out in Table 3; as the tonogenesis process often involves a spread of laryngeal features into neighboring phonetic environments I am using "environment" here to mean the *phonological* environment where the triggering feature is specified.

Table 3: Domains of *de novo* tonogenesis

Environment	Level tone generated	Contour tone generated
Onset	Lhasa Tibetan (Hsieh & Kenstowicz 2008)	???
	Punjabi (Ohala 1978)	
	New Caledonia (Rivierre 1993)	
Nucleus	Hup (Class I, see 4.2; Epps 2008)	Hup (Class II, see 4.2; Epps 2008)
	Utsat (L, see 4.3; Thurgood 1999)	Utsat (LH, see 4.3; Thurgood 1999)
		Vietnamese (huyền; Thurgood 2008)
Rime	Utsat (HL, see 4.3; Thurgood 1999)	Vietnamese (nặng; Thurgood 2008)
Coda	Utsat (H, see 4.3; Thurgood 1999)	Khaling (Mazaudon 1977:65-6)

3 – Phonology of tone

With the phonetics relevant to tonogenesis introduced, we now turn our attention to theories of how tone is represented in languages' phonology, and why tonogenesis poses an interesting problem for certain representations of contour tones.

3.1 – What is phonological tone?

"Lexical tone" was defined in 2.2 as a feature of certain languages whereby lexical items are distinguished by tone. This is done in different languages to varying degrees. On one end of the spectrum is Cantonese, where every syllable of every lexical entry is specified for tone. Mandarin has some environments where syllables do not bear lexical tone, and thus there is not only a contrast between tonemes, but between lexical tone and zero. For example, /tɕî-tɤ̌/ "vested" where the second syllable is specified for a LH tone contrasts with /tɕî-tɤ/ "remember" where the second syllable lacks lexical tone and is assigned tone post-lexically. Reduplicated material is also stripped of its tone and reassigned tone post-lexically, for example /ʂɨ̂-ʂɨ̂/ "human affairs" is composed of two homophonous morphemes and contrasts with /ʂɨ̂-RED/ "try" where the reduplicated /ʂî/ does not retain its HL tone and is assigned tone post-lexically.[9] Other languages such as Hup (Nadahup, Brazil; discussed in-depth in 4.2) contrast tones only on stressed syllables, the unaccented syllables either having arisen through epenthesis (representing no material in the lexicon and thus no lexical tone) or being high-frequency unstressed function words (cf. Mandarin toneless /tɤ/ "GEN" /mɤn/ "AN.PL"). While some phonologists may argue that languages with environment-restricted tonal contrasts might not be "truly tonal," some compromise must be made. No phonologist would be willing to say that Mandarin is not a "true tonal language" despite the fact that some Mandarin morphemes lack lexical tone. As I am interested in how tonal information enters the lexicon, for my analysis I will consider all languages whose lexicon contains tonal information.

"Pitch accent" is a phenomenon in certain languages that is analyzed by some authors as lexical tone. Pitch accent systems are described by Hyman 2006:236 as exhibiting some aspects of stress systems and some aspects of tone systems. Stress systems are those that have a hierarchy of prominence: certain syllables of lexical items will be specified for this prominence, or "stress." Pitch accent systems may be analyzed as having a hierarchy of syllable prominence that surfaces in prominent syllables having a certain tone (while other stress systems may use other signals for this contrast such as amplitude, or a combination of the two). For example, Blevins 1993 describes Lithuanian pitch accent, a system in which accented morae surface with an H tone (Blevins 1993:239).

The commonality of pitch accent systems with both stress and tone has spawned analyses of these systems through the frameworks of both stress and tone. Analysis of pitch accent as tone is straightforward: tones are associated to certain syllables/morae in the lexicon. Under Blevins

[9] Examples from Harbaugh 1999.

1993’s analysis of Lithuanian, H tones are underlyingly associated to the morae and remain in the surface representation, while unaccented morae are assigned tone post-lexically through the prosody.

The alternative stress-based analysis of pitch accent rejects the notion of tones in the lexicon and instead places a “diacritic” (*) denoting prominence. This diacritic then acts as an anchor to which tonal “melodies” are strung. Pulleyblank 1986:155 demonstrates this with an example from Luganda, [àbàpákàsì] “porters.” The surfacing LLHLL tone melody suggests the third syllable as prominent, giving it a diacritic in the underlying representation. An LH*L melody then anchors to the prominent syllable, forming the intermediate representation [abàpákàsi]. The first and fifth syllables then gain an L tone through spreading. The advantage of this approach in this example is that it identifies the third syllable as the one that “stands out.” However, a purely tonal analysis can also set the third syllable apart by having it the only one specified for tone in the lexicon, with the L tones assigned by a default rule (Pulleyblank 1986:160). Pulleyblank uses such examples to question the necessity of the diacritic as a phonological primitive, deriving its supposed benefits through a purely tonal framework.

Whether the information in these cases specified in the lexicon is tonal or some form of prominence is not a trivial ambiguity. The latter analyses imply that the pitch accent systems they describe are not lexical tone. As this paper intends to explore the process by which languages acquire lexical tone, languages that do not exhibit lexical tone cannot be considered, and a cutoff for what is and isn’t a “tonal language” must be defined.

Hyman 2006 attempts to establish clear definitions of stress, tone and pitch accent systems. He is unable to establish any prototype of a pitch accent system, instead relying on negative definitions (of the type “a pitch accent system must lack some phenomenon X of stress/tone systems") (Hyman 2006:236). A true cut-off point between pitch accent systems and stress/tone is also not possible as pitch accent systems share different sets of commonalities with stress/tone systems, and these commonalities lack any hierarchy that would make adherent languages more or less "tonal" (Hyman 2006:237-8).

For the sake of this paper I must consider "pitch accent" languages to be tonal. Many of these languages use F_0 as a primary perceptual cue, for example Japanese has pitch accent, but the only significant marker of "accented" syllables is fundamental frequency, while stress is not well-evidenced (Pierrehumbert & Beckman 1988:7). Another reason to consider pitch accent languages as tonal is that they are able to distinguish lexical items in isolation purely through pitch. An example from Japanese is [kàkí] "persimmon" vs. [kákì] "oyster" (Matsumura 1988).[10]

The most relevant commonality between pitch accent and tone is that languages often develop pitch accent through the reinterpretation of some other phonetic feature that affected F_0, precisely the phenomenon I seek to discuss. Examples of tonogenesis in pitch accent systems will be discussed in 4.1.

3.2 – Synchronic consonant-tone interaction

Consonant-tone interaction was discussed in 2.4. To recap: consonants often have phonetic features that affect F_0. This makes them obvious candidates for triggers of tonogenesis, and

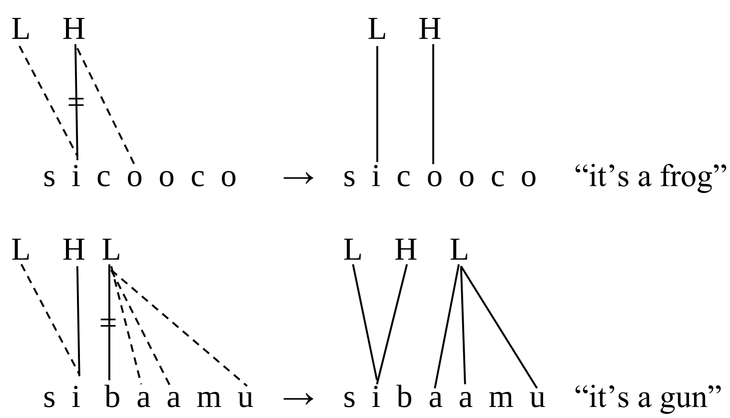
[10] Complicating the issue, the Japanese pitch accent system is such that minimal pairs can also be distinguished from tonal information that has spread past the original lexical item. For example, [kákì-ò] "oyster-ACC" and [kákì-ó] "fence-ACC" can be distinguished by the tone of the following postposition. While these may resemble "melodies" that have been anchored to these lexical items, Pierrehumbert & Beckman 1988 derive the Japanese tonal pattern without the use of diacritics, instead employing lexical tone in combination with boundary tones. This tonal analysis does not require sets of arbitrary melodies with no discernible place in synchronic systems (Pulleyblank 1986). As pitch accent is traced back to Proto-Japanese without much further internal reconstruction (see Ramsay 1979), a discussion of Japanese in the context of tonogenesis is not possible in this paper.

numerous examples of this were listed in 2.5. However, consonant-tone interaction is not limited to diachrony. 3.1 led a discussion on what it means for tonal information to be present in the lexicon, choosing to use such an analysis whenever tone is the primary cue by which surfacing material is contrasted.

Tone in the lexicon as discussed in 3.1 typically takes the form of autosegments associated to the syllable. However, tone can also be a feature specified to consonants. The most common occurrence of this is the “depressor consonants” widely attested in African languages. An example is Siswati (Nguni, South Africa) where the voiced obstruents /b d g ǁ̬ v z ɮ d͜v d͜z ɦ dʒ/ are always followed by either an L or LH tone. For example, *kúvùúka* “wake up” but **kúvúúka* is impossible (Bradshaw 1999:11-12). These depressor consonants also block the shift of H tones. An example of this is the Siswati predicative construction: an H tone on the initial syllable is shifted to the penult and an L tone is associated to the first syllable. The presence of a depressor consonant in the second syllable onset blocks H shift, and LH surfaces on the initial syllable. Examples:

(1) a.

b.



In (1a), the H from the first syllable shifts to the first mora of the penult and the grammatical L tone is associated to the first syllable. The depressor consonant *b* in (1b) blocks H from moving to the penult, leaving it trapped in the first syllable. The L also associates to the first syllable, forming an LH contour, while the L tone from the depressor consonant spreads over the last two syllables.[11]

While the “canonical” depressor consonant is a voiced consonant with an associated L tone, H tones are also available for association to consonants, and the consonants to which the tone is associated are not necessarily the natural class of voiced consonants. Tang 2008 does a more thorough typology of these, sorting them by “affinity for L” and “affinity for H,” meaning that they participate in some sort of consonant-tone interaction or distributional restriction that suggests specification of that tone. Among the languages surveyed, voiced consonants and slack/lax consonants have an affinity for L tones, stiff/tense consonants have an affinity for H and voiceless consonants, implosives, ejectives, sonorants, aspirates, fricatives, and glottal consonants all have at least one instance of being specified for each tone (Tang 2008:25-6). For example, just as the depressor consonant in (1b) blocked spreading of a H tone, in Bade (Chadic, Nigeria) L-tone spreading is blocked by voiceless consonants (Schuh 2002).

Moreton 2009 compares these synchronic consonant-tone interactions with synchronic interactions between phonological tones. He finds that tone-tone interactions are much more frequent than consonant-tone interactions within languages with lexical tone. Such a finding cannot be explained by phonetic salience alone as consonants have a greater effect on tone than

[11] Bradshaw 1999 does not discuss the fate of the syllables unmarked for tone in *sìcóoco*, nor does she specify a discrete underlying/surface representation. Presumably this is just one process within a greater tonal phonology in Siswati, and I have left out mention of morpheme composition and levels of representation to be faithful to the data. Bradshaw also uses her data to deduce that depressor consonants such as the one in (1b) are specified not for L but for a feature that encompasses both privative voicing and low tone. However, this analysis has been dismissed as overly simplistic, for example by Tang 2008.

other neighboring tones (Moreton 2009:2). While Moreton's survey is imprecise,[12] his greater conclusion that languages tend towards interactions between like elements (consonants/vowels/tones) is interesting. Moreton explains this result through the constraint-based Optimality Theory framework: phonologization of an interaction between like elements would only require a reranking of the constraints relevant to that type of element while phonologization of dissimilar elements requires reranking of the constraints relevant to both initial elements. Moreton calls this "modularity bias," constraints governing consonants and tones can form discrete modules with no inter-module rankings until necessitated by some

[12] Moreton's survey only considers one language per family as a "precaution against...shared inheritance or areal spread," however this backfires in five ways. First, it discounts cases of interaction within language families that are not shared inheritance (his example in Indo-European is Chakma (East Indo-Aryan), which excludes unrelated tone-tone interactions in Scandinavian languages, for example). Second, the ability to spread to unrelated languages *defines* areal features. Moreton's inclusion of Sino-Tibetan, Tai-Kadai, Hmong-Mien (and in later discussion Mon-Khmer) shows that contact/areal spread of tonal features has not been ruled out. A third problem is that languages may share other features which then bias them towards tone-tone or consonant-tone interactions. The obvious example is depressor consonants in African languages, where longer word length and affixation allow for phenomena of tone shift, which is one of the major indicators of depressor consonants and is not supported by the mostly-monosyllabic tonal languages of Asia. The fourth issue is that an "accidental gap" in consonant-tone interaction can stem simply from the loss of features such as voicing during tonogenesis. While Moreton tries to compensate for this by considering only languages that have some VOT contrast that affects tone, it cannot be ignored that VOT contrasts in Asia are largely in aspiration and not voicing as is found in African languages. While both aspiration and voicing affect pitch, they do not do so by the same mechanism (see Section 2), and the difference in ability to control tone or maintain tonal contrasts in environments with voicing/aspiration may bias the presence or absence of consonant-tone interaction. Finally, by discounting languages without contrastive lexical tone, Moreton ignores languages like Hindi, where tone is predictable (but more exaggerated than natural consonant-tone interaction, see Ohala 1978) and governed by some sort of phonology but not contrastive. In order for these predictable tones to surface correctly under OT, the relevant constraints must be reranked even though tone is not contrastive. As Moreton uses his survey to generate statistics, an accurate result must consider all languages that have undergone this reranking. Unfortunately due to the areal nature of tone, no survey of the world's tonal languages will truly avoid the influence of areal spread and perfectly accurate statistics are impossible. However, with no reason to limit such a survey by language family, a much larger sample size can be used for more accurate results.

phonological consonant-tone interaction (and likewise for other modules such as the vowel-related constraints). Because of this bias, Moreton characterizes interactions that necessitate inter-modular rankings more costly than those that don't. While this may be the case under a constraint-based framework that privileges efficiency, it is also important to discuss the implications of Moreton's typological findings under other theories of phonology.[13] These results will be revisited in 5.5 in the context of phonologization and teleology.

3.3 – Tone and Distinctive Feature Theory

There are multiple proposals for how lexical tone fits into Distinctive Feature Theory. One matter of contention is the representation of phonological contour tones. There are two main proposals for feature-based contour tones. The first, which I will call the Contour Tone Features (CTF) analysis, posits that contour tones occupy a single feature matrix. This feature matrix will either include specifications for the presence of a tone differential (something like [±CONTOUR] contour for zero vs. non-zero tone differentials) or for its values ([RISING] for a positive tone differential value and [FALLING] for negative values, or a binary feature that encompasses them) or both. An example of this proposal is from Wang 1967, who additionally posits separate features for non-static tone differentials (convex and concave):

[13] While the difference between constraint-based models of phonology like Optimality Theory and rule-based models are relevant to phonologization (such as Moreton's findings), my main focus is the representation of inventories, not interactions. For this, I assume Distinctive Feature Theory and Autosegmental Phonology, both also assumed by Optimality Theory.

Figure 3: feature chart following Wang 1967

	˥	˩	˦	˨	˧	˧˥	˩˧	˥˧	˧˩	˥˧˥	˧˩˧	˧˥˧	˩˧˩
[CONTOUR]	-	-	-	-	-	+	+	+	+	+	+	+	+
[HIGH]	+	-	+	-	-	+	-	+	-	+	-	+	-
[CENTRAL]	-	-	+	+	+	0	0	0	0	0	0	0	0
[MID]	0	0	-	-	+	0	0	0	0	0	0	0	0
[RISING]	0	0	0	0	0	+	+	-	-	+	+	+	+
[FALLING]	0	0	0	0	0	-	-	+	+	+	+	+	+
[CONVEX]	0	0	0	0	0	0	0	0	0	-	-	+	+

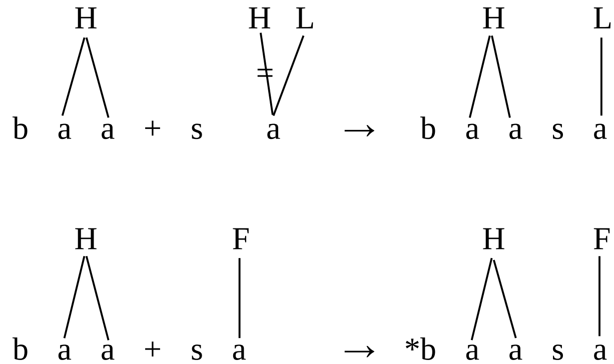
The other approach to phonological contour tones I will call the Autosegmental Contour Tones (ACT) analysis, where contour tones are divided into multiple level tone values which occupy discrete tone-bearing units (TBU's) within a syllable. For example, a falling tone would be represented as a sequence of high tone + low tone (HL), assigned to the same syllable.

Evidence for the ACT analysis is found in morphological processes where single TBU's within a contour tone are phonologically active, or where tones move according to changes in syllable structure. Goldsmith 1976 (the dissertation that formalized autosegmental phonology) gives many such examples, one of which is the deletion of level tones within contours via the Obligatory Contour Principle (OCP; a constraint that prohibits adjacent identical tones/elements). From Goldsmith 1976:138:

49A) /báá/ 3pl.PAST + /sâ/ "grind" → [báásà] *Ganda (Bantu, Uganda)*

Here, the underlying form /sâ/ has a discrete H and L tone associated with one TBU each. As the H in /báá/ is identical to this first H on /sâ/, it deletes leaving [sà]. A CTF analysis would have some featural specification for “falling” in the matrix for /sâ/; without the high tone feature this would not trigger the OCP.[14] Illustrated for each analysis:

(2) a. b.



(2a) uses an autosegmental representation of the falling tone on /sâ/, with distinct H and L tones in the HL falling contour. When the inflectional prefix is added, this H tone is now adjacent to the H tone in /báá/, and deletes to satisfy the OCP. (2b) instead has a unitary falling tone F on /sâ/. As this representation does not contain an H tone, the H in /báá/ does not cause an OCP violation and the result is the unattested *[báásâ].

Another piece of evidence for surfacing contour tones having an ACT-style representation is floating tones, or morphemes with tonal information not associated to any segment. An example from Yip 2002:88:

(3) /ip1/ “Yip” + /a3- -5/ (hon.) → [a3 ip15] “Miss Yip” *Cantonese*

[14] Note that the CTF analysis merely proposes that features specifying the tone differential exist, not that all surfacing contour tones are governed by single feature matrices with these features. My mention of “the ACT analysis” is not the idea that autosegmental phonology exists in general, but that autosegmental contours can account for all surfacing contour tones. For more discussion see Yip 1989, 2002, Gussenhoven 2004, and Clements & Patin 2011. In addition, while the OCP is claimed to be exceptionless only morpheme-internally, it is also used to describe language-specific cases such as in (2) where like adjacent elements are prohibited across a morpheme boundary. The OCP is thus used to describe (2) following the original example in Goldsmith 1976.

In (3), when the title /a3/ is added to /ip1/, a floating /5/ tone must be added to the end of the name. (Or /a3 _ 5/ could be analyzed as a circumfix.) This floating /5/ toneme is productive with all names and not part of /ip1/'s underlying representation. Thus, in the surfacing [ip15], the [1] and [5] must be separate phonological entities, level tones in separate TBU's.

African languages often have both tone and active morphology, and some show alternations between level and contour tones when the morphology affects the number of syllables. Newman 1995:1.1 lists examples from Hausa where verbs and pronouns can optionally be mono- or disyllabic but retain tone:

(4) a. [zân ~ záánì] "I will"

b. [mîn ~ mínì] "to me"

There are other aspects of tone that may be analyzed in multiple ways. Another theory in tonal phonology regards *tone registers*, where a feature like [±HIGH REGISTER] specifies whether the tone is in the top half of the tone space while another (perhaps [±HIGH TONE]) describes the tone's position within the register. In these systems, languages with four or five contrasting level tones can be distinguished, for example San Andrés Chicahuaxtla Trique with four level tones (Hollenbach 1977:50, Yip 2002:214-6) or dialects of Black Miao and Tahua Yao with five (Miao-Yao; Chang 1953:375, Yip 2002:27). Tone registers are also useful for languages that contrast contours with the same tone differential in different parts of the space; this can be seen in Vietnamese tonal harmony, where reduplicated material unspecified for tone is assigned a level tone according to the register of the source material's tone (Nhàn 1992). Tone register features can also be used for processes where a tone is raised over another tone, for example Clements 1978 describes a process in Ewe where mid tones (M) are raised to an extra-tone higher than the language's existing H tone. Here, the M tone is specified as

[-HIGH REGISTER +HIGH TONE], and its shift to [+HIGH REGISTER] raises it higher in the tone register than H, which is [+HIGH REGISTER -HIGH TONE].

Another type of feature system for tones uses articulatory features to bridge tone with consonantal features, for example using [SLACK VOCAL FOLDS] for low tone and consonant voicing (Stevens 2000:251, Laver 1994, Bradshaw 1999). Feature systems like these derive synchronic phenomena like depressor consonants, (usually) voiced obstruents that are specified for L or act as a barrier to tonal movement (Hyman & Schuh 1974:105-6). While these proposals are relevant to accounts of tonogenesis, they are simply too simplistic to capture certain distinctions (such as tone vs. phonation type) that are essential to the analyses in this paper.

Additional proposals may eschew tone features altogether (Clements et al. 2011), however this paper will assume at least a distinction in tone height (e.g. H, M, L), and that this distinction involves some feature(s) in the phonology regardless of whether they fully satisfy the goals of Distinctive Feature Theory. For example, it is unclear from the discussion in Clements et al. 2011 what proposed mechanism would account for interactions in tonal and prosodic phonology without a featural distinction for tone height.

3.4 – Un ou deux tonèmes?

Another issue in tonal phonology is how to distinguish whether a non-zero tone differential in the surfacing phonetics acts like a level or contour tone in the phonology. A well-known example is the Mandarin "third tone," which acts as a single L tone in the phonology (and surfaces as L in many prosodic environments) but in elicitation is pronounced with a [214] concave tonal contour (Yip 1980, 2002). While few tones surface with a completely level F_0, tonal phonologists generally distinguish level from contour tones in the phonology depending on how they behave. For example, Yoruba (Kwa, Nigeria) has three contrasting tones, phonetically

[45 33 31], however in the absence of evidence to the contrary they are treated as phonologically level H, M and L tones (Gandour 1978:48).

Evidence of tones behaving phonologically as contour tones under the ACT analysis include the examples in 3.3 of individual TBU's within a contour engaging in phonological interactions. As the CTF analysis proposes additional features, evidence of a phonetic contour tone acting like a contour phonologically would involve the proposed features being referenced in phonological processes.

The problem of whether a contour in the surfacing phonetics is underlyingly one complex autosegment or multiple phonological entities mirrors that of complex segments such as affricates and diphthongs, often called the "un ou deux phonèmes" problem after Martinet 1939. Single-segment affricates can contrast in a language with separate stop-fricative sequences, for example Polish speakers can distinguish *trzy* [tʃɨ] "three" with a two-segment onset from *czy* [t͡ʃɨ] "whether" with one complex segment (Martinet 1939:98). Hayes 2009:56 similarly presents a near-minimal pair in English between the tautosyllabic diphthong in *boing* [bɔ͡ɪŋ] and the sequence in *sawing* [sɔɪŋ] which surfaces heterosyllabically.[15] Contour tones can also act as both units and multiple phonological entities within a language: Chan 1991 describes a tone-spreading process in Danyang Chinese in which either the latter half of a contour tone or the entire contour tone spreads. For example, an underlying /HL L/ will spread over a four-syllable domain as

[15] Hayes' suggestion that the diphthong in *boing* is a complex segment is controversial. However, at the least, such an analysis for the heterosyllabic diphthong in *sawing* is impossible. Further evidence would be needed to characterize the former as a single complex segment, while the latter must be two segments under any analysis.

[HL L L L] but an underlying /HL LH/ will spread over a four-syllable domain as [HL HL HL LH] (Chan 1991:241).[16]

Splittability is also good evidence that an element with two phonetic targets is two phonological entities. An example is the resyllabification of Dutch diphthongs: if roots ending in diphthongs [iu eu ui oi ai] precede a vowel-initial suffix, the latter target becomes a glide [w j] in the onset of the suffixed syllable, for example *fraai* [frai] "beautiful" becomes *fraai-e* [fra.jə] "beautiful (attributive)". (Van Wijk 1939a, Zonneveld & Trommelen 2002:277) Dutch also has the diphthongs [ɛi œy ɔu], however these do not undergo the same process, for example *kei* [kɛi] "boulder" but *kei-en* "boulders" is [kɛi.ə], not *[kɛ.jə] (Zonneveld & Trommelen 2002:277). This evidence points to the diphthongs [ɛi œy ɔu] being single phonological units while the splittable [iu eu ui oi ai] are two segments.[17] Splittability in tone is shown in the Hausa example in 3.3, where an alternation in number of syllables causes an alternation between one contour or two level tones.

Other kinds of evidence that are useful in distinguishing a multiple-target phone as a single vs. multiple phonological entities are found in phonotactics and typology. Hayes 2009:56 notes that many languages with diphthongs only have certain licit vowel sequences, while other permutations of the inventory's monophthongs are unattested as sequences. This overprediction is solved by having the attested diphthongs in the phonemic inventory. Yip 2002:50 provides a similar argument for tonal inventories, that many languages with contour tones do not exhibit

[16] It is assumed from the spreading of HL towards the end of the domain in the /HL LH/ case that spreading in the /HL L/ case is in the same direction and not simply spreading of L in the opposite direction. Cases where the first tone is a contour with a second target not matching the second tone are absent as Danyang Chinese is restricted to six "tone patterns" (Chan 1991:239).
[17] What exactly the underlying representations of [ɛi œy ɔu] depends on the analysis. Van Wijk 1939b:253 proposes a system in which [ɛi œy ɔu] are underlyingly close-mid vowels /ɛ œ ɔ/ with offglides added in the phonology. For further discussion, see Zonneveld & Trommelen 2002.

every possible sequence of the level tones in the inventory, thus strengthening analyses of tone systems where contour tones are units present in the phonemic inventory.

Having one of the two targets in a complex segment/tone not otherwise present in the phoneme inventory has also been used to tip the scale towards unitary segments/tones. Dutch has the diphthong [œy] but no monophthongal *[œ], mentioned by Zonneveld & Trommelen in their discussion of [œy] as an underlying monophthong (see Footnote 17). Yip 2002:60 describes the Mandarin rising tone as [35], but Mandarin does not otherwise have [3].[18]

Speaker intuitions can also be taken into consideration when analyzing multiple-target tones/segments as single or multiple phonological entities. Pike 1947 describes the set of diphthongs in American English as [aɪ ɔɪ aʊ eɪ ɪi oʊ ʊu]; he was able to get naïve native speakers to separate the targets in [aɪ ɔɪ aʊ] but not in [eɪ ɪi oʊ ʊu], suggesting that speakers intuit the latter set as monophonemic and the former as multiple phonological entities. Intuitions about tones as single/multiple phonological entities are less often discussed in the literature, however some perceptual experiments have courted the idea. Abramson 1976 collected data on the perception of the five tones in Thai (/H M L HL LH/): the M tone surfaces with a fall near the end, and speakers considered a level [33] tone "abnormal" for /M/ but would categorize it as /M/ as easily as they did the [332] contour that Thai speakers produce (Abramson 1976:122). The Thai speakers were thus cognizant of the small fall at the end of the M tone, but a tone differential was not necessary for the categorization of that tone. Perhaps this suggests that Thai

[18] Yip 2002 ultimately adopts an ACT analysis of contour tones. This data was mentioned in the context of evidence for the CTF model, and the author herself agrees that neither ACT nor CTF are fully satisfactory representations of tone in phonology.

speakers consider M a level tone phonologically, even if they are aware that it is not level phonetically.[19]

3.5 – Predictions of contour tone analyses

In 3.4, the problem was addressed of how to tell whether a surfacing non-zero tone differential was acting as one or multiple phonological entities. Clearly, if a contour tone exhibits behavior such as splittability, it lends itself to an ACT-style autosegmental contour. However, of the contour tones described in 3.4 as behaving as a single phonological entity, it must be decided what that phonological entity is. It could underlyingly be a level tone to which a second target is added post-lexically, for example the /M/ tone in Thai surfacing as [332]. It could be a complex autosegment with an intermediate level containing multiple phonological tones. This type of representation has been posited for Danyang Chinese: Chan 1991:243 represents the HL tone as an H and an L tone within a node between the syllable and tone tiers; this node is then copied allowing for the full HL fall to spread onto following syllables. Having multiple phonological tones below this node (and not a contour governed by CTF) is useful for other kinds of spreading in Danyang Chinese, where only one tone in the contour spreads.

Of course, the "single phonological entity" contour tones in 3.4 could simply be single feature matrices governed by contour tone features. As the CTF analysis does not prevent multiple feature matrices from forming an autosegmental tone contour (only positing that

[19] Additionally, Chilin Shih (personal communication) described the Mandarin "third tone" (phonologically /L/, phonetically [214]) as level, but in need of an on- and off-glide to "dip" down to the low level. Clearly this is not an actual articulatory need as [11] tones are common cross-linguistically. As Shih is not a naïve speaker and there are sandhi rules that give better credence to /L/ as a phonological level tone, I have not included a discussion of her speaker intuitions. However, it seems that some sense of "level-ness" or "contour-ness" (whether that correlates to single vs. multiple phonological entitie(s) or information about the tone differential) is accessible to speakers. Further research is necessary to fully analyze these kinds of speaker intuitions.

features exist such that a contour tone *can* be governed by a single matrix), it is difficult to completely rule out the CTF model from the above evidence. Instead, we must look to predictions of another sort: predictions made by the feature content of CTF-only features.

The main difference between the CTF and ACT analyses in terms of feature content is that CTF's features encode properties of the tone differential. Take a falling tone (HL) for example. According to the ACT analysis, this tone is represented in the phonology as two level tone, H + L. The gradient fall in F_0 that surfaces is thus something that arises in the phonetics, the natural coarticulations between adjacent tones differing in level. Neither an H tone nor an L tone contain any featural information about the tone differential; under the ACT analysis the resulting tone differential is governed by the linear order of the individual tones and their encoded tone heights.

In order to get to the heart of the featural content that governs tone, it is important to ask what properties of the pitch contour are perceived by speakers of a language on the cusp of tonogenesis. Do they perceive a series of pitch height values? Or can they directly perceive the tone differential? If certain properties of tone are perceived by these speakers, it is these properties that should be encoded in post-tonogenetic synchronic grammars.

It may be intuitive to associate perception of the tone differential with the CTF analysis. If there is some environment in which F_0 rises (i.e. has a positive tone differential value), translating this perception to a featural specification for [RISING] is seamless.

The role of the tone differential in ACT analyses is less apparent. As the featural information encoded in tones under the ACT analysis is purely that of tone height, one must explain the prevalence of inconsistencies in this tone height information in the synchrony and diachrony. For example, in Utsat (Chamic, China), an H tone is described phonetically as [5] when alone, but as [4] when in a contour HL [42] or LH [24] (Maddieson & Pang 1993:77). From the survey of tonal inventories in Yip 2002 it is clear that this is not an isolated case. But is it really necessary to posit a phonological process for languages of this type by which tones are centralized in the pitch range in the presence of an opposing tone, always within the domain of the syllable (despite the widespread tonal interactions over syllable boundaries, see Moreton 2009 and Goldsmith 1973)? Or are [4] and [5] separate tonemes restricted to complementary environments? The intuitive solution, that the information in the lexicon that effects falling tones in the surface representation is some variation of "falling," requires that the tone differential be encoded in some fashion; under the ACT analysis this fashion must be indirect.

3.6 – The interface between phonetics and phonology in tonogenesis

The path from phonetic environments with a shift in F_0 to contrastive sequences of discrete level tones in the phonology is not straightforward, as it is for the CTF analysis which can encode the tone differential directly. The most glaring oddity in an ACT analysis (to this author) is that multiple phonological tones result from a phonetic environment that could be characterized by a single phonetic feature.[20]

This paper seeks to address the discrepancy between the predominant ACT theory of contour tone phonology and phonetics-based accounts of tonogenesis. Are there environments that necessitate single vs. complex tonogenesis? Must they arise simultaneously, or can the discrete tones in an autosegmental contour be the result of multiple stages of tonogenesis

[20] Proponents of ACT may argue that this single phonetic feature (i.e. one referring to the tone differential) does not exist and that rises/falls would always be categorized by the listener into discrete levels (preserving a 1:1 ratio in perceived- to phonologized features). While features of the tone differential can be argued against as features available for contrasting lexical items, speakers must be able to perceive rises/falls in F_0 for independent reasons, such as prosody and the perception of glottal stops. For a discussion of when and how rises/falls in F_0 are perceived, see Gandour 1978.

(telescoping)? Do the resulting autosegmental tone sequences match their triggering environments in tone height, or can they diverge in tone height but share a tone differential value?

It seems that while ACT is the dominant theory in regards to representing contour tones in the phonology, its proof lies largely in the synchrony. In the following sections I will discuss various accounts of the evolution of contour tones, and attempt to use this data to bridge the phonetics of tonogenesis to the phonology of ACT-style tone.

4 – Case studies

In 3.2, cases of consonant-tone interaction were discussed in which synchronic phonological processes resulted in surfacing contour tones. An example was Siswati, where voiced consonants are specified with L, creating an LH contour when in the onset of a syllable with high tone. This section will instead examine the *diachronic* origins of contour tones through various case studies in order to bridge synchronic tonal systems to the tonogenesis processes by which they arose. 4.1 will describe contour tones entering the lexicon in pitch accent systems that have shifted in syllable structure. In 4.2 we will see Hup (Nadahup, Brazil), where contour tones have arisen through telescoped processes from a non-tonal origin. 4.3 will conclude the section with the Chamic languages and Utsat in particular where a full tonal inventory has arisen without telescoping from a non-tonal origin.

4.1 – Contour tones evolving from non-lexical level tones

The question of what counts as tonal information in the lexicon was addressed in 3.1. Prosody also makes use of tones, and these are uncontroversially not part of the lexicon. These tones, however, can be reinterpreted on lexical items. This often creates the partially-tonal pitch accent systems, where tone melodies from the prosody are reinterpreted on certain lexical items (as opposed to something like a tone split triggered by laryngeal features that affects all lexical items).

As prosody uses “melodies” of multiple tones, it is no surprise that reinterpretation of prosody can result in the phonologization of contour tones. For this reason I have included two examples of pitch accent systems in which the reinterpretation of pitch from prosody has generated contrasts that resemble those of “true tonal languages” if there is such a thing. In these examples, single-syllable lexical items can be distinguished by F_0 contours alone, in the first example the contrast is between two contours tones and in the second is between a level and a contour tone.

The first example is Scottish Gaelic (Ladefoged et al. 1998, Iosad (to appear)). Intervocalic glides such as /w w̃/ (orthographically <bh, mh>) dropped and their flanking vowels fused, merging with existing diphthongs. As Gaelic has initial stress, this left an early pitch peak on the fused words which contrasts with the later peak in other words. An example from Ladefoged et al. 1998:

(5)	<dubhan> “hook”	*[ˈtu.wan]	>	[túàn]
	<duan> “song”	*[ˈtuan]	>	[tùán]

The second example is Franconian German, which deleted word-final schwas (Gussenhoven 2000, 2004, 2013, Gussenhoven & Peters 2004). These schwas were unstressed and had a low F_0, reinterpreted as a low tone that was then reassigned to the surviving previous syllable. An example from Gussenhoven 2013:3:

(6)	*dax “day” (c.f. Standard German *Tag*)	>	[dááx]
	*ˈda.ɣə “days” (c.f. Standard German *Tage*)	>	[dáàx]

Because these tones arise from the reassignment of existing tonal patterns, they are not examples of tonogenesis. However, they show an evolution of contour tones that fits in well with the ACT theory of contour tones: the units within the contour specify only tone height and originate from discrete phonetic environments.

It must be noted that the phonologization of prosodic melodies is not restricted to the syllable as in (5-6). However, when the affected domain is greater than the syllable, it is difficult to determine where and how many tones have been phonologized, or which tones are lexical and which are governed by post-lexical rules. As the representational discrepancy between the ACT and CTF models addressed by this paper involves contour tones within the syllable, I have chosen the above examples to contrast contour tones within this domain. Additionally, the examples are such that the contrastive tonal information is *only* within the domain of the original lexical item, unlike examples such as Japanese where contrastive tonal information extends to clitics and outside the prosodic word through what resembles a "melody" (see 3.1, and Footnote 10).

4.2 – Hup

Hup is an example of a language with contour tones that did not evolve from the reassignment of existing level tones. In this case, the units within the autosegmental tone contour originate from discrete (phonologically atonal) environments at different points in time.

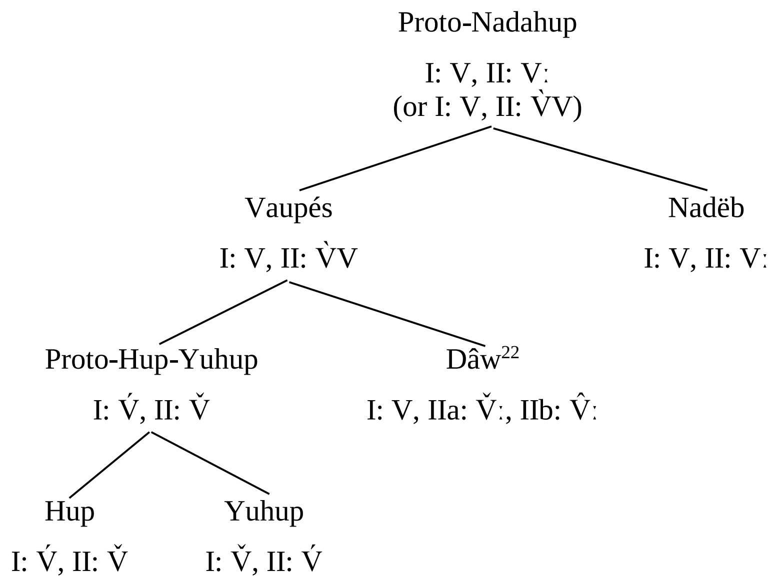
4.2.1 – The Nadahup languages

The Nadahup languages are a family in northwestern Brazil and areas of Colombia and Venezuela. These languages have two classes of words: one (which I will refer to as Class I[21])

[21] Epps 2008 uses the term "noun classes" to refer to noun genders. Since that information is not relevant to this analysis I will use "Class I/II" as a shortcut for the set of cognates sharing

has short vowel length in Nadëb and Dâw, high tone in Hup and rising tone in Yuhup. The second (Class II) has long vowels in Nadëb, rising tone in Hup, high tone in Yuhup and a contour tone in Dâw depending on its coda. The family tree with reflexes of the noun classes is shown in Figure 4 (Epps 2008, Patience Epps personal communication):

Figure 4: Evolution of word classes in the Nadahup languages



While it's not certain whether Proto-Nadahup had a length distinction which became tone in the Vaupés languages or a tonal distinction that was reinterpreted as vowel length in Nadëb (Patience Epps, personal communication), some suprasegmental feature in Proto-Nadahup governs two sets of reflexes in the four modern Nadahup languages. Hup and Yuhup have a degree of mutual intelligibility and share 90% cognate basic vocabulary; however Yuhup has undergone some form of tone reversal as Yuhup Class I stems behave phonologically like the Class II stems in Hup and Dâw and the converse (Epps 2008:3).

---

suprasegmental features as shown in Figure 4. These classes include nouns, adjectives and possibly verbs (Epps 2008:88).

[22] Dâw Class II nouns underwent a split; those with voiceless codas became IIa and those with voiced codas became IIb (Patience Epps, personal communication).

4.2.2 – Tonal phonology in Hup

Hup morphemes are of the structure (CV)CVC with the two vowels of identical quality if present.[23] In addition, only the vowel with stress bears lexical tone (high or rising). As there are also no licit onset/coda clusters in Hup, this suggests an underlying representation of /CCVC/ for surfacing disyllabic stems.[24] These clusters would be broken up by an epenthesized vowel copy, and the references to morpheme boundaries in the phonology (Epps 2008:50-72) could be replaced with reference to the syllable boundary. While Epps (personal communication) notes that certain synchronic compounding processes might pose a problem for such an analysis, such an underlying representation is at least very likely in an earlier stage of the language.

The high-toned Class I stems are subject to an additional phonological process where they become falling tones if the stem has a voiced or nil coda, as seen in (7) (Epps 2008:89,91):

(7)	a.	/tə́g/ “tooth”	→	[tə̂g$^{\text{ŋ}}$][25]
	b.	/j’á/ “black”	→	[čâ̰ː]
	c.	/tĕ̌g/ “wood, stick”	→	[tĕ̌g$^{\text{ŋ}}$]

In (7a), the stem’s high tone surfaces as a falling tone because of its voiced coda, and the nil coda in (7b) causes the same effect. (7c) shows a Class II stem which does not undergo the same change in tone.

---

[23] Counterexamples exist but generally are loanwords or can be traced to compounds, e.g. /~yʔáb.hŏ/ “dog” from /~yʔáb/ “jaguar.” (Epps 2008:84) The tilde ~ marks nasality which is a feature of the syllable in Hup. Examples in this section are written in the original Americanist notation.

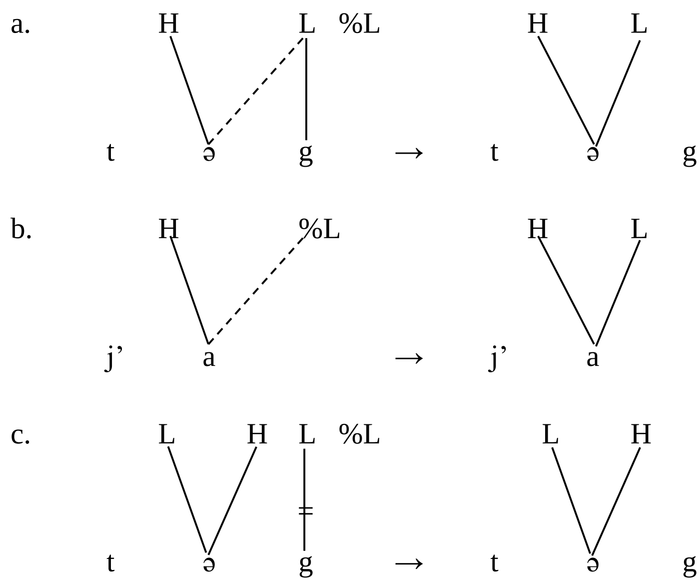
[24] While some initial-stress words exist and would be analyzed as /CVCC/, their status as native Hup words is debated. (Epps 2008:87)

[25] Nasal release/prenasalization surfaces when an underlying voiced consonant is present at a morpheme boundary.

4.2.3 – Telescoping in tonogenesis

With the data in 4.2.1 and 4.2.2 we can begin to unpack the evolution of Hup tones. Synchronically, it seems that any syllable that can bear lexical tone has two TBU's: Class I stems have an H on one TBU and an empty second TBU which can accept an L associated with a depressor consonant or the %L boundary tone. Class II stems are unaffected as both TBU's are full with their lexical LH rise. This is shown in (8a-c), corresponding to the examples in (7a-c):

(8)



Epps also notes that the Class I stems in Dâw are atonal, while the class IIa/b stems have a rising/falling intonation. These contour tones are not necessarily specified for tone on both TBU's; they could easily have a tone assigned to the first TBU and the second is either filled in by the prosody or tones associated with the coda consonant. Such an analysis would look much like the above analysis of Hup without the H tones (Class I unspecified for tone and Class II with an L tone on one TBU and a second empty TBU). For this reason I have marked the Class I/II contrast in Vaupés (and possibly Proto-Nadahup) as V vs. V̀V in Figure 4.

The H tones in Hup occur on all stressed syllables and on no unstressed syllable. Is it a coincidence that these unstressed syllables all resemble vowel copies? If not, it could be that the contrast between these epenthesized vowels and the (yet toneless) Class I nouns was maintained by associating an H tone with all stressed syllables.

Yuhup's "swapped" tone classes suggest some sort of tone reversal in its development. While this tone reversal could be recent (after developing a tonal system like modern Hup), there are differences between the Hup-Yuhup reversed word classes and other cases of tone reversals that may indicate an early reversal. Athabaskan (Kingston 2005) is a well-known case of tone reversal: one group of languages phonologized an H tone onto a certain class of stems ("high-marked languages"), while the cognate stems in other languages in the family have an L tone ("low-marked languages"). For example, "father" is *tó* in low-marked Navajo but *tù* in high-marked Chipewyan (Kingston 2005:168). Kingston proposes several pathways from pre-tonogenesis Proto-Athabaskan to the modern languages with opposing tones: for example, as the "marked" class of stems originally had glottalization, the opposing reflex tones in high- and low-marked languages could have arisen due to the realization of this glottalization as tense/creaky voice respectively (see 2.3.2) (Kingston 2005:163). This tone reversal involves variations in articulation with different effects on pitch that are then phonologized as opposing tone levels, creating an "upside-down tones" effect. Such an analysis doesn't work in Nadahup, where one level- and one contour tone, each with their own phonology, seem to have been "swapped." A more likely scenario is that the reversal happened before the shared innovation of epenthesis/assigning H tones to stressed syllables. Yuhup marked Class I stems with L instead of the Class II stems; after the shared innovations with Hup (including the depressor consonants and %L boundary tone), this resulted in an identical tonal phonology to Hup's but with the stem classes reversed.

It seems that the contour tones found in all three Vaupés languages have arisen due to multiple stages of tonogenesis. The level tones that make up the autosegmental contours originate from discrete processes during different periods of the languages' history: first from an earlier contrast in length, then from prosody and contrast with new epenthesized vowels and finally from an interaction with segmental features and boundary tones. These processes are shown for Hup in Table 4:

Table 4: Possible evolution of tones in Hup

Syllable structure and class	I: CVT	I: CVD	I: CCVT	I: CV	II: CVC
Example stem, original form	*~tæh	*təg	*btɔk	*j'a	*teeg
L-tone on long	-	-	-	-	*tèeg
Length distinction neutralized	*~tææh	*təəg	*btɔɔk	*j'aa	*tèeg
Epenthesis	-	-	*bɔtɔɔk	-	-
H-tone on stressed syllable	*~tǽh	*tə́g	*bɔtɔ́k	*j'á	*tèég
Modern Hup UR	/~tǽh/	/tə́g/	/b(ɔ)tɔ́k/	/j'á/	/tĕg/
L-tone from depressor consonants	-	tə́ə̀g	-	-	(deleted)
Boundary %L tone	-	-	-	j'áà	-
SR (after other phonology)	[tǽ̃h̃]	[têg$^{\text{ŋ}}$]	[^{m}bɔtɔ́k]	[čậː]	[tĕg$^{\text{ŋ}}$]
Tone of surface form	H	HL	H	HL	LH
Gloss	"son"	"tooth"	"ear"	"black"	"wood"

4.3 – Chamic

The Chamic languages are a sub-branch of Malayo-Polynesian within the Austronesian language family, spoken in western Indonesia, Vietnam and surrounding areas. The Chamic languages underwent language contact with the tonal languages in the region, and in some cases had radical changes in syllable structure as well as genesis of register and tone. Utsat, a Chamic language spoken on Hainan Island, China, underwent the most radical shift, from initial CVCVC disyllable roots to single CVC syllables which contrast for three level tones H, M, L and two contour tones LH, HL.

Chamic in particular is a useful account of tonogenesis as it can be reconstructed with great detail. Thurgood 1993, 1999 (as well as Maddieson & Pang 1993) provides an account of historical Chamic (and particularly Utsat) that shows the role of phonation in creating the environments that trigger certain kinds of tonogenesis. This case study will show an example of contour tones arising without telescoping from a single phonetic environment.

4.3.1 – Syllable restructuring and register in Chamic

The structure of the Proto-Chamic (PC) root was $C_1VC_2VC_3$; henceforth references to indexed segments such as C_2 or V_1 will be in reference to the order of C/V's in the PC root (e.g. C_3 for the coda of the second root syllable in PC).

One of the first changes in Chamic was the genesis of breathy voice after voiced consonants, likely due to the shared low F_0 of voiced obstruents and breathy voice; the voiced consonants became voiceless (aspirated in Utsat). This breathy voice spread towards the end of the disyllabic root, in Utsat over sonorants and voiceless obstruents other than [h]. (Thurgood 1993:98) In Phan Rang Cham (PRC; Thurgood 1993) and Western Cham (WC; Thurgood 2008), this same spreading occurred over [h] but not the other voiceless obstruents curiously (Thurgood 1993:94).

V → V̤ / D X _ Utsat: X = T, (D), N, but not [h]

WC/PRC: X = (D), N, [h] (and *s > [h])

This breathy voiced register is maintained in WC, and surfaces as a low tone in PRC main syllables.[26] After phonation-spreading, Utsat underwent a reduction in syllable structure. V_1 deleted, reducing the root structure to CCVC with an onset cluster formed by C_1 and C_2. These clusters were reduced; if C_2 was a liquid [l r] it became a palatal glide and then [i]. In other cases, C_1 was deleted. The phonation spreading and syllable reduction can be seen in the following examples from Thurgood 2008:[27]

(9)

		PC root	Phonation spreading	Utsat σR	Utsat IF	WC reflex	Gloss
a.	i.	*ribɔw	*ribɔ̤w	*(r)bɔ̤w	*phə̤		"thousand"
	ii.		*ribɔ̤w	-	-	ripɔ̤w	
b.	i.	*jalaan	*ja̤la̤a̤n	*(j)la̤a̤n	*la̤a̤n		"path"
	ii.		*ja̤la̤a̤n	-	-	ca̤la̤n	
c.	i.	*bituk	*bi̤tṳk	*(b)tṳk	*tṳʔ		"cough"
	ii.		*bi̤tuk	-	-	pətuʔ	
d.	i.	*jahit	*ja̤hit	*sit	*siʔ		"sew"
	ii.		*ja̤hi̤t	-	-	chi̤ʔ	
e.	i.	*lima	-	*(l)ma	*ma		"five"
	ii.		-	-	-	lamɨ	
f.	i.	*labuh	*labṳh	*(l)bṳh	*phṳh		"fall down"
	ii.		*labṳh	-	-	lapṳh	

[26] Chamic languages that maintain disyllabic root structure do so either as iambs or sesquisyllables. The term "main" syllable refers to the second syllable, or the reflex of $C_2V_2C_3$ generally.

[27] Intermediate forms are mine; Thurgood describes them but does not show them in the reconstruction tables. "Utsat IF" is the intermediate form of Utsat after the processes described in this section. They are a landing point from which the discussion in 4.3.2 picks up. "Utsat σR" is the syllable restructuring process undergone in Utsat but not WC. Parenthesized segments in the Utsat σR column are either deleted C_1 or C_2 glides surfacing as [i]. Some phonological changes (including *jh > *s, *T > ʔ / _# in Utsat) appear in the reconstruction but are not relevant to the discussion on register and tone. <j> in these reconstructions is a voiced palatal obstruent; <y> is the palatal glide.

In (9), the (i) examples show the derivations in Utsat and (ii) those in WC. In the (a,f) examples, C_2 is voiced while C_1 is not; breathy voice is spread onto the main syllable as can be seen in the WC reflex. (9b) shows a case where C_1 is voiced and C_2 is a sonorant allowing for spreading into the main syllable in both Utsat and WC. (9c,d) show examples where C_2 inhibited spreading in one of the languages; in (c) the voiceless [t] prevents spreading to the main syllable in WC but not Utsat while in (d) the [h] prevents spreading to the main syllable in Utsat but not WC. (e) is an example where neither C_1 nor C_2 is voiced; both Utsat and WC reflexes are modal voiced.

4.3.2 – Utsat tonogenesis

The "Utsat IF" column in (9) is the stage in historical Utsat is the stage of the reconstruction containing all information relevant to tonogenesis. From the reconstructions in this column we can track the origins of Utsat tones.

Modern Utsat has three level tones (H, M, L) and two contours (HL, LH). The contour tones arose on syllables with a coda glottal stop (to which PC coda stops had neutralized); a rising tone LH on the modal register or a falling tone HL on the breathy register. Of the syllables without coda glottal stops, those with coda [h] received a high tone H, breathy voiced syllables without coda glottal consonants gained a low tone L and the remaining modal voiced syllables without coda glottal consonants phonologized a mid tone M. This can be seen in the following examples:

(10)

PC	Utsat IF	Modern Utsat	Tone	Gloss
*labuh	*phṳh	phú	H	"fall down
*lima	*ma	mā	M	"five"
*ribɔw	*phə̤	phə̀	L	"thousand"
*jahit	*siʔ	sǐʔ	LH	"sew"
*bituk	*tṳʔ	tûʔ	HL	"cough"

Now let us focus on the phonetic environments that triggered the genesis of these tones. The level tones are straightforward: the low F_0 of the breathy voice register is replaced with a low tone, the [h] is replaced with a high tone (see the discussion in 2.4.1) and in the absence of other phonetic triggers, the modal voiced syllables are phonologized with a mid tone to contrast with the other tones.

The motivations for phonologization of the contour tones requires more analysis. It is interesting that while both environments (modal voice + glottal stop for the rising tones and breathy voice + glottal stop for the falling tones) share a coda glottal stop, their resulting tones share only that they are contours. Under an ACT analysis, what is shared by these tones is only that they are each composed of two phonological tones. But how do these tones arise?

One option would be for the tones to arise from the nucleus and coda (a one-to-one correspondence between segments and phonological tones). However, this idea does not fit the data. If the glottal stop is associated with a fixed tone height (say, M) and the nucleus gains a tone based on its phonation (L for breathy voice and H for modal voice), the contours would have the opposite tone differentials from what we see in Utsat (LM for breathy + glottal stop is rising where Utsat has HL falling). A one-to-one correspondence between phonological tones and nucleus/coda (or segments) in the triggering environment is thus not possible.
Thurgood 2008 (and Gage 1985) describes a situation similar to the Utsat rising tones in Vietnamese, where rising tones were associated with syllables that had some glottal element in the coda. Glottal stop codas are often realized as some [CONSTRICTED GLOTTIS] phonation on the latter part of the vowel (see the discussion in 2.4.2); if this was the case then the two phonological tones that make up the rising contour in Utsat could have arisen from the two

halves of the vowel. As the tone differential is positive, this suggests that the [CONSTRICTED GLOTTIS] phonation in question is high-F_0 tense voice and not low-F_0 creaky voice. An example of the process:[28]

(11)

			L H		
	CVʔ	→	CVV̰ʔ	→	CV̌ʔ
Ex. *jahit “sew”	*siʔ		*sìíʔ		sǐʔ

The breathy voiced syllables that preceded the Utsat falling tones differ from their modal counterparts in one crucial way: unlike modal voice, breathy voice is incompatible with other phonations like tense voice (Laver 1980:132). If the vowel was already fully breathy, tense phonation cannot bleed into the nucleus. The path from breathy vowel + glottal stop without a stage of [CONSTRICTED GLOTTIS] phonation in between is most easily understood with reference to the tone differential. Breathy voice is already quite low in terms of tone height, and it is not immediately obvious how the breathy voice vowel would be assigned an H tone (as part of the HL contour in Utsat falling tones). However recalling the experimental data discussed in 2.4.2, it seems that the best cue for perception of glottal stops is a *dip* in F_0. Even if the breathy voiced vowel is already low in pitch height, the glottal stop causes F_0 to fall even lower. This creates a

[28] The IPA does not have distinct diacritics for creaky vs. tense voice; I will use V̰ to designate tense voice in this section. In addition, the notation “VV̰” to designate a vowel that begins in modal voice and ends in tense voice does not designate that the vowel has two phonological segments or timing units (this is impossible as all stages from PC to Utsat have maintained a distinction in vowel length). Instead this can be considered a “phonation contour,” as described for example in DiCanio 2009 for the Takhian Thong Chong breathy→tense register.

negative tone differential, expressed in the phonology as a series of tones that fall in pitch, i.e. a falling tone. Here is an example of the process:[29]

(12)

			L XL		H L		
	CV̤ʔ	→	CV̤ʔ	→	CV̤ʔ	→	CV̂ʔ
Ex. *bituk "cough"	*tṳʔ		*tùṳ̏ʔ		* túṳ̀ʔ		tûʔ

Note that the coda glottal stops have behaved in two different ways in the development of Utsat. As the vowels in the breathy-voiced words are already specified for phonation, the realization of this glottal stop as tense voice on the latter half of the vowel is blocked. As the tense-voiced vowel and true glottal stop have opposing effects on pitch, they triggered the genesis of opposing tones.

5 – A bridge from tonogenesis to autosegmental contour tones

In earlier sections we have discussed the phonetics and phonology of tone as well as the phonetics of tonogenesis. This section will attempt to use the case studies in Section 4 to bridge the phonetics of tonogenesis to the autosegmental model of contour tones and address the questions posed in 3.6.

5.1 – An evolutionary framework of phonologization

In the literature, most theories of phonology describe synchronic grammars while work dealing with the diachrony is often examined through a purely phonetic lens. The most notable exception

[29] In this example I use XL as an "extra-low" tone, as it is a tone introduced into the system that is lower than the existing L tone in the system. This notation is often used in the Africanist literature, see Yip 2002.

to this is the Evolutionary Phonology framework, which seeks to formalize sound change and relate sound change to trends in sound patterns of the world's languages (Blevins 2004). Blevins' typology of sound change identifies three main pathways, as described in the table below (from Blevins 2004:32-8):

Table 5: Typology of sound change under Evolutionary Phonology

CHANGE		Perceptual similarities between phones A and B cause [A] to be misheard by listeners and subsequently reinterpreted as /B/.
	Ex.	Speaker intends /f/, says [f]. Listener hears the phonetically-similar [θ] and interprets /θ/.
CHANCE		The speech stream is heard correctly, but the listener interprets the surface representation as the result of an underlying representation and phonology different from the speaker's.
	Ex.	Speaker intends /aʔ/, says [ʔa̰ʔ]. Listener hears [ʔa̰ʔ] correctly, but interprets an underlying /aʔ/ where the laryngealization spreads from a glottal stop coda instead of a glottal stop onset.
CHOICE		An single phonological form has multiple phonetic outputs. The listener hears these variants and chooses an exemplar from them, different from the speaker's underlying representation, interpreting it as their own.
	Ex.	Speaker intends /kakata/, produces variant set [kakáta], [kkáta], [kăkáta]. Listener chooses [kkáta] as the exemplar, interpreting the underlying representation as /kkata/.

In order to fit tonogenesis to these models, we must first determine how tone relates to the pathways in Table 5. Many instances of tonogenesis involve environments with phonetic features that affect F_0. Both the change in F_0 and the original phonetic feature that caused it are phonetic cues in the environment; tonogenesis from this perspective is the choice of tone as the primary cue and the reinterpretation of underlying forms to be specified for tone instead of the other phonetic feature. Thus, CHANCE is the most important of the three pathways for tonogenesis.

An (abstracted) example of CHANCE:

(13)	Speaker		Listener
	/da/		/tà/
	↓1		↑3
	[d̥à...]$_{Vi}$	2→	[d̥à...]$_{Vj}$

In (13), the speaker intends /da/ with a voiced onset and no specified tone. In production (1), this onset is realized with some of the articulatory gestures that facilitate obstruent voicing (lowering the tone of the syllable) but with too wide a glottis to support voicing during the stop closure, surfacing in [d̥à] (Laver 1994:343-4). The listener hears the signal correctly (2), but interprets the surface representation as the result of tone on the lexical item instead of voicing (3).

CHANGE and CHOICE also play a role in tonogenesis. Features that lower F_0 often also have variants with creaky voice, for example. This allows for the listener to choose an exemplar (CHOICE), resulting in the speaker's utterance being reinterpreted either as the original feature (no change), tone, phonation, or tone and phonation (like in the Vietnamese register complex). An example of CHANGE might be something like Proto-Athabaskan dialects switching between tense/creaky voice due to their phonetic similarity (see 2.3.3). While not tonogenesis itself, this change played a pivotal role in the development of Athabaskan's languages with opposing lexical tones.

In order to generalize about the featural content of tones (and thus the featural content of contour tones), I must note the following relationship:

(14) Speaker Listener

$/[X]_{\delta}^{[A]}/$ $/[X]_{\delta}^{[B]}/$

$\downarrow 1$ $\uparrow 3$

$[[X]_{\delta}^{[C]}]_{Vi}$ $2\rightarrow$ $[[X]_{\delta}^{[C]}]_{Vj}$

(14) is a CHANCE model, where domain δ (a syllable, foot, mora...) contains segmental material X. A and B are phonological featural material on the suprasegmental tier while C is some phonetic environment that can be analyzed as the result of either A or B plus phonological rules. The speaker's underlying representation contains A but is reinterpreted as B through the ambiguity of C.

The possible content of A and B is the object of discussion in Distinctive Feature Theory, as discussed in 3.3. Under a CTF analysis, A and B can contain direct information about the tone differential, while under the ACT model, A and B can only be contours as some complex autosegment such as $[A_1\ A_2]_A$.

What C is allowed to contain is another issue. The content of A and B must be grounded in some feature of C, otherwise listeners could phonologize features irrelevant to the perceived features of an utterance. For example, hearing the [+ROUND] in [to] and phonologizing it as a nasal /tã/. Clearly this is not a process common to the world's languages. Thus, whether we deem certain content legitimate in A and B affects whether we deem this content a "perceivable" feature of C. With this in mind, let us turn our attention to the data surveyed in section 4.

5.2 – Sources of phonological contour tones

The ACT model of contour tones posits that a contour tone consists phonologically of a series of level tones. Contour tones that exist in the world's languages today have multiple origins; they can be innovations from tones active outside of the lexicon (as in prosody), or they can be the result of tonogenesis.

Within the cases of contour tones arising from tonogenesis there are the examples like Hup where the discrete tones that constitute the autosegmental contour were phonologized at different points in the language's history, and there are cases like Utsat where a single instance of tonogenesis in a certain environment generated multiple phonological tones. Not only is it possible under the ACT analysis of contour tones for more than one of these tones to arise simultaneously, but these simultaneous tones can originate from the same segment/timing unit, for example the Utsat rising tone described in 4.3.2.

In addition to the phonetic/phonological environments that lend themselves to tonogenesis, language contact is another source of phonological tone. Tone is notorious as an aerial feature: languages that come in contact with tonal languages have a much higher chance of developing lexical tone themselves. (Blevins, to appear) Utsat and Vietnamese are both languages from traditionally atonal families; both acquired large tonal inventories after a period of heavy contact with Chinese. The Scots Gaelic dialects that developed pitch accent did so after contact with Swedish, which has a similar pitch accent system (Riad 1996, 1998, 2006).

It must be asked whether if there is any difference in the phonologization of level vs. contour tones under contact. While this may be the case, there is an important intermediate step between contact and tonogenesis that may bias tonogenesis towards certain kinds of tones. There is non-tonal information that spreads through contact such as phonation and prosody. Since these features are not contrastive in many of the languages that undergo tonogenesis under contact, these languages are free to adopt them in order to gain some sociolinguistic benefit (or alternatively, gain them subconsciously in a non-teleological manner). If these features in turn predispose languages to certain paths of tonogenesis, there could be such a predisposition towards phonologizing level or contour tones.

5.3 – The phonologization of single vs. multiple tones

We must also ask why a tonogenesis process in general would result in a contour of multiple phonological tones instead of a single one. For example, why did Utsat result in three level tones and two contour tones instead of five level tones? As noted in the discussion in 3.3, a system with five contrastive level tones is possible. If the Utsat tone system stemmed from a contrast between five phonological environments, why did some of these become level tones and others contours?

There are a number of ways to approach this question. One is that of maintaining contrast. 4- and 5-level tone systems generally have contour tones in addition to the numerous level tones (e.g. the Miao, Yao and Trique varieties noted in 3.3). Perhaps a tonal inventory with 4-5 level tones is more stable when the language also contains contrasting contour tones. If features of the tone differential are perceptible and able to be phonologized (i.e. they can be in C in the discussion in 5.1), Utsat may have preserved contrast by phonologizing contour tones where the original environments had tone differentials farthest from zero, forming a stable tonal inventory.

Another possibility is that there are certain phonetic features that provoke the phonologization of multiple tones over single ones. For example, the dip in F_0 perceived in glottal stops (see 2.4.2 and the Utsat falling tone in 4.3.2) is *relative* to the preceding vowel; the experimental data show that the negative tone differential of glottal stops is a significant perceptual cue while tone height is not. If the primary phonetic cue is a property of the tone differential, it makes sense for a listeners to reinterpret this as some phonological representation with a tone differential value.

5.4 – The fate of the tone differential

Recalling the discussion of Evolutionary Phonology in 5.1, we revisit the model of suprasegmental sound change:

(14) Speaker — Listener

$/[X]_{\delta}^{[A]}/$		$/[X]_{\delta}^{[B]}/$
$\downarrow 1$		$\uparrow 3$
$[[X]_{\delta}^{[C]}]_{Vi}$	$2\rightarrow$	$[[X]_{\delta}^{[C]}]_{Vj}$

In (14), the allowable phonological feature content in A and B is determined by the set of distinctive features (however they are analyzed). Under the CTF analysis, features that encode information about the tone differential exist and are available for inclusion in A and B. ACT analyses reject this: the only tonal information allowed in A and B are features of tone height (and register if applicable). As the perceptible phonetic features in C must match those available in A and B (see discussion in 5.1), proponents of the ACT model would prefer for features of the tone differential to not be acceptable in C as well. While ACT has succeeded in excising features of the tone differential from A and B, can observed accounts of tonogenesis be explained without these features being available for C? Let us amend the structure in (14) to show an autosegmental contour being phonologized:

(15) Speaker — Listener

$/[X]_{\delta}^{[F]}/$		$/[X]_{\delta}^{[K1\ K2]}/$
$\downarrow 1$		$\uparrow 3$
$[[X]_{\delta}^{[G1\ G2]}]_{Vi}$	$2\rightarrow$	$[[X]_{\delta}^{[G1\ G2]}]_{Vj}$

In (15), some suprasegmental featural content F surfaces in a phonetic environment with two distinct values of some perceived phonetic feature G: $[G1\ G2]_G$. This is heard correctly by the listener who then interprets this environment as the phonological contour $[K1\ K2]_K$. What is the relationship between the values of G and the values of H? Does G1 have to match K1 and G2 match K2, the way that C had to match B in (14)?

Such an analysis does not hold up for examples like the Utsat falling tone. In this case, G is pitch (and K is tone height): G1 is the pitch caused by breathy voice and G2 is the “dipped” pitch of a glottal stop. If G1 were phonologized to K1, this would correspond to the breathy-voiced vowel phonologizing to the H in the HL falling contour tone. However, breathy-voiced vowels without glottal stops (G1 without G2) phonologized to L in Utsat, not H.

Appeals to Structure Preservation will also not solve the issue. Phonetically, breathy voice has a low F_0, and the expected phonologization of breathy voice is an L tone. The reason breathy voice phonologized to H (G1 to K1) in the Utsat falling tone is not because the breathy voiced vowel was somehow closer to the existing H-tone category, but because it had a higher F_0 relative to the glottal stop.

The only elegant solution in this case is to adopt the (14)-style model with features of the tone differential available for C. For the Utsat falling tone, C is a negative tone differential value, and is phonologized to B which is some featural content like [FALLING]. Such an analysis derives the commonality between the original triggering environment ([21]) and the phonologized tone (/42/), that is to say the tone differential value.

So then how can the need for representation of the tone differential be satisfied under the ACT model? This problem may be a problem of autosegmental phonology in general, as “sequence-dependent” autosegmental representations exist outside of tone. Recalling the $[K1\ K2]_K$ representation in (15), this mold has been used in autosegmental phonology for complex segments such as affricates (where K is [CONTINUANT]). The ability to represent affricates as [-CONTINUANT +CONTINUANT] allows feature systems to shed features that only contrast affricates (such as [±DELAYED RELEASE]). However, such a representation does not derive the fact that affricates are robust in the phonemic inventories of the world’s languages but the equally-admissible [+CONTINUANT –CONTINUANT] contour is not. This asymmetry is “sequence-dependent”, it is information that references the sequence of featural information in addition to its content. In the example of the Utsat falling tone, the negative tone differential that was referenced in phonologization requires that the two targets (K1 and K2) be in the correct order and have values relative to each other. Autosegmental phonology has no way of encoding such a sequence or relative value in its framework.

This is not to say that autosegmental phonology or the ACT model of contour tones is incorrect. Even in cases such as Utsat, the triggering environment had two distinct targets. Evidence remains for single targets within contour tones being referenced by the phonology (for example the Danyang Chinese tone spreading). In order to reconcile the lack of encoding sequence and relativity in autosegmental phonology, some additional kinds of representation are necessary, such as coindexation. In the representation $[K1\ K2]_K$, the indices of the targets within the contour could be referenced by some sort of phonology or feature. For example, features such as a negative tone differential could be represented in some method such as $[K1 > K2]_K$ or with some phonological rule that can reference the indices of the targets within the contour. In this way, features of the tone differential could exist not as simply a feature in a matrix without hierarchy but some formula dependent on a feature that dominates it in the feature geometry.[30]

A representation like this of relativity and sequence in autosegmental phonology would not only benefit a representation of contour tones, but other sequence-dependent phonology such as sonority in syllabification.

5.5 – The teleology of tonogenesis

One last question remains to be addressed in this paper: why tonogenesis at all? Over half of languages do not exhibit contrasting lexical tones.[31] Even languages that are in contact with tonal prestige languages have managed to not acquire tone, such as Mongolian and Finnish. The idea of tonogenesis as a "trade-off" has been humored by many authors, such as Matisoff 1973 (Maddieson 2007:94): this idea implies that the redundancy of phonetic environments contrasted by both tone and some other phonetic feature (such as voicing) provokes neutralization of this redundant feature when tone begins to be contrastive.

This model of tonogenesis as a "trade-off" is not entirely accurate, however. Languages may start to contrast environments by tone but retain a second contrasting feature, for example

---

[30] The quantification of feature values in a representation such as $[K1 > K2]_K$ also poses a problem for the largely-held view that features are binary. Even within this, a "+" featural specification could be seen as existing on some continuum with "-", making comparisons such as "greater than" meaningful. However, such a view of binary features is probably not faithful to the proponents of binary feature theories, who may prefer other concepts like "markedness." A solution for the hierarchy of feature values is outside the scope of this paper. It must be noted, however, that these phonological features reference phonetic qualities that can be quantified (and compared) such as F_0.

[31] According to WALS, which notes that the ratio of tonal languages may be greater if the Niger-Congo languages were better represented.

the glottal stop codas in Utsat made redundant by the contour tones, or the predictable tones in Hindi in environments contrasted primarily by laryngeal onset features (Ohala 1979).

Maddieson 2007 addresses the "trade-off" idea statistically, looking for a correlation between the richness of a language's tonal inventories and simplicity in other features like consonant and vowel inventory size. He finds a positive correlation between the size of tonal inventories and the sizes of consonant and tonal inventories, and concludes that as these consonant/vowel inventories are as rich as the tonal inventories that no trade-off has occurred between them. He also finds a slight negative correlation between size of tonal inventory and complexity of syllable structure. Between this, and Maddieson not taking into account language features such as the size of roots (number of syllables, etc.) and the type of morphology (isolating or otherwise), it seems that the door is still open for tonogenesis to have some teleological value, or at least for tonogenesis to facilitate other changes in languages such as the simplification of syllable type and presence or absence of affix-based morphology. Another consideration is the sociolinguistic benefit of approximating one's speech patterns to those of a prestige speech community under contact.

According to the ACT analysis, a syllable with a contour tone contains more information than a syllable with a level tone (as opposed to each having one CTF feature matrix with different feature values). If phonologists are looking at tonogenesis from the perspective of proving or disproving an "equilibrium" of information that is maintained by exchanging phonetic features for lexical tone, autosegmental models of contour tones predict that they will have a different effect than level tones on this process.

The idea of "amounts" of phonological information and how they affect phonologization recalls the discussion of Moreton 2009 in 3.2. Moreton's finding was that interactions between like elements phonologize more frequently than equally-salient interactions between dissimilar elements. As tonogenesis from laryngeal features is essentially "consonant-tone interaction" and happens much less frequently outside of contact situations than other kinds of phonologization, it is worth examining whether the infrequency of *de novo* tonogenesis from laryngeal features can be accounted for by a Moreton-style "modularity bias."

Moreton's explanation under Optimality Theory was that interactions between like elements were more easily phonologized as they require reranking of only the constraints relevant to one type of element while interactions between dissimilar elements require reranking of the constraint modules that govern each type of element as well as rankings between constrains in these previously-discrete modules. Another justification for the costliness of phonologizing consonant-tone interactions involved OT's view that certain "marked" phonetic environments are somehow problematic, and that the constraint rankings in the phonology somehow solve these problem environments. For example, the "marked" environment *LH only has one solution, spreading of L to surface in [LL] (Moreton 2009:9).[32] However, in situations with dissimilar elements (such as a consonant-tone interaction), multiple solutions are available:

[32] While the reasoning behind "markedness" of specific tone height values is not discussed by Moreton, his model assumes that H is marked and L unmarked (as does Yip 2002). Thus [HH] is not a solution as it creates a more marked environment than the input. Interestingly, this approach is opposite of Bradshaw 1999, whose collapsed feature "voice/L" assumes that voicing and L would share one (presumably marked) feature specification. Both theories seek to account for the relationship between obstruent voicing and L tones; while Bradshaw 1999 proposes that the two share a quality, OT describes the same data as the two features having opposing values of another property (markedness). The asterisk (*) here denotes a marked phonetic environment (and the phonological constraint that prohibits it), not ungrammaticality, reconstruction or affinity with stress.

the “marked” environment *bH (voiced consonant adjacent to a high tone) can be fixed either by devoicing the consonant to [pH] or by lowering the tone to match the voiced consonant as [bL]. As *bH has two less-marked fixes under this OT model, Moreton posits that it is an interaction that is less likely to be phonologized (Moreton 2009:12).

If tonogenesis from laryngeal features suffers some “modularity bias,” we must find some equivalent environment to Moreton’s *bH where both consonantal and tonal modules must undergo reranking. However, there are some key differences between relevant tonogenetic environments and *bH. Under Moreton’s model, the difference between tone-tone and consonant-tone interaction stems from the presence of two dissimilar *phonological* entities, however the initial state of languages that undergo tonogenesis from consonant features is unspecified for tone. The surfacing environments of the type [pL] in languages like post-tonogenesis Chinese and Vietnamese are not the fix of an earlier [bH].[33] They are the result of an ambiguous stage that could be the result of either consonantal or tonal phonology. For example, environments like [dà] surface in both Hindi and Lhasa Tibetan. In Hindi, the low tone stems from the voicing of the onset; the contrast between [dà] and [tá] stems from an underlying contrast in voicing (/da/ vs. /ta/). However, the same [dà]-[tá] contrast in Lhasa Tibetan is the result of an underlying contrast in tone (/tà/ vs. /tá/) (Tournadre & Dorje 2003:32-4). This is the nature of CHANCE: two equally valid phonologies could result in a surfacing sound pattern, and a listener chooses a different one of these phonologies from the speaker.

If the infrequency of tonogenesis outside of contact situations stems from Moreton’s “modularity bias” (i.e. the nature of interactions between dissimilar elements and the additional

[33] In fact, under Moreton’s teleological model, [pL] would exhibit *both* proposed fixes of [bH] where both [b] and H have lost their marked features to satisfy *bH.

parameters involved in their phonologization), then phonologization of [dà] as /tà/ must be more costly than phonologization of [dà] as /da/. Each of these paths only requires reranking within a single module, however. The only way for the /tà/-type phonology to be more costly than the /da/-type phonology is if this choice is made by a listener who already has the necessary phonology for consonant laryngeal features but does not have an active tone "module." But are these the conditions under which tonogenesis occurs? In order to answer this, we must determine two things. First, does tonogenesis occur more easily in first-language acquisition, where /da/- and /tà/-type phonologies are equally costly without interference from pre-existing phonology? (And do other types of phonologization work differently?) Second, can whatever "module" of phonology that governs lexical tone be activated in speakers by other phonology such as prosody (erasing the extra cost of /tà/-type phonologization)?

It seems that if rates of phonologization can be described as evidence of inherent phonological "biases," the infrequency of *de novo* tonogenesis from consonant features cannot be explained by the same bias as Moreton describes for synchronic consonant-tone interaction. In addition, some bias against phonologizing lexical tone as a whole would not derive the bias described by Moreton (as all the languages in his survey had previously phonologized lexical tone). Clearly, much additional evidence is needed to make claims that the phonologization of tone is affected by some psychological "cost" or teleology. Instead, if there is some bias that accounts for the dearth of consonant-based tonogenesis and consonant-tone interaction, it would more likely be a bias in where listeners map F_0 cues. If they are more likely to hear F_0 cues as a redundant feature of the previous consonant than as a separate suprasegmental feature, the low rate at which both lexical tone and consonant-tone interactions are phonologized would be derived from this bias.[34] However, this too would require evidence.

To extend this reasoning to contour vs. level tones, an ACT analysis of contour tones without the additional representations of tone differential information (as proposed earlier in this section) would not have any additional "cost" in phonologization compared to level tones. On the other hand, both the CTF model and representation of the tone differential in an autosegmental framework (as discussed in 5.4) would require access to phonology outside of whatever "module" governs level tones. Under these analyses, contour tones would be less easily phonologized due to Moreton's modularity bias. While any claims of modularity bias in contour tones would be premature, it has been noted in studies such as Clumeck 1980 and Ma et al. 2012 that Mandarin-speaking children acquire the (high) level tone more quickly than the contour tones (the last of which is always the concave [214] tone).

6 – Conclusion

While proponents of synchronic theories of phonology seek to shed structures and features that they deem redundant, speakers do not only reference features and phonological representations

[34] Blevins 2004:152 describes the *feature-to-segment mapping principle* (FMP) whereby listeners interpret a feature F on some longer domain as a feature on some source segment /S_F/. A bias towards a tonal feature being heard on a consonantal source approximates this model, however the case of tone is special as what is phonologized by listeners who do not assume a tonal autosegment is often not a tonal feature on the consonantal source /C_F/ but a completely different feature on the source /C_G/ where G is some laryngeal feature that is realized as phonetic information that includes F. As the FMP itself does not suppose any hierarchy between the domains onto which F is phonologized, I have restricted discussion of the biases towards/against these domains to Moreton's teleological model. However, if these biases do exist, then the low rates of phonologization of tone and consonant-tone interaction would be a result of this mapping process.

in the synchrony. A “workaround” for some structure that makes a theory of phonology less efficient may gloss over the properties of that structure referenced during its phonologization.

The dominant theory of contour tones rejects any representation of features of the tone differential. However, properties of the tone differential not only are perceived by speakers but shape the phonologization of tone. Thus, in this paper I have argued that autosegmental phonology needs some way of representing properties of the tone differential, through sequencing of the targets within contours and the ability to reference this sequencing in the phonology. With this ability, autosegmental phonology can account for the process of phonologization in tonogenesis, as well as derive asymmetries in segmental phonology.

Bibliography

Abramson, Arthur S. (1976). "Static and dynamic acoustic cues in distinctive tones" in *Haskins Laboratories Status Report on Speech Research*, SR-47, pp.121-127.

Anderson, Stephen R. (1978). "Tone Features" in *Tone: A Linguistic Survey*. Ed. Victoria A. Fromkin. Academic Press. pp.133-176.

Backley, Philip and Kuniya Nasukawa (2013). "The role of L in the pitch accent system of Tokyo Japanese" in *Phonological Studies*, Vol. 16.

Beddor, Patrice S. (2009). "A Coarticulatory Path to Sound Change" in *Language*, Vol. 85, No. 4 (Dec. 2009), pp. 785-821.

Blevins, Juliette (2004). *Evolutionary Phonology*. Cambridge University Press.

Blevins, Juliette (2007). "Interpreting Misperception: Beauty is in the Ear of the Beholder" in *Experimental Approaches to Phonology*, Ed. M-J. Solé et al. pp. 144-154. Oxford University Press.

Blevins, Juliette (to appear). "Areal Sound Patterns: From Perceptual Magnets to Stone Soup" in *The Cambridge Handbook of Areal Linguistics*. Ed. Raymond Hickey. Cambridge University Press.

Bradshaw, Joel (1998). "Another Look at Velar Lenition and Tonogenesis in Jabêm" *Oceanic Linguistics*, Vol. 37, No. 1 (Jun. 1998), pp.178-181. University of Hawai'i Press.

Bradshaw, Mary (1999). *A Crosslinguistic Study of Consonant-Tone Interaction*. Diss. Ph.D. Ohio State University.

Catford, John C. (1977). *Fundamental Problems in Phonetics*. Indiana University Press.

Chan, Marjorie K. M. (1991). "Contour-Tone Spreading and Tone Sandhi in Danyang Chinese" *Phonology*, Vol. 8, No. 2, pp.237-259. Cambridge University Press.

Chao, Yuan Ren (1930). "A System of Tone-Letters" in *La Maître Phonetique*, Vol. 45, pp. 24-47.

Chang, Kun (1953). "On the Tone System of the Miao-Yao Languages" in *Language*, Vol. 29, No. 3 (Jul-Sep 1953), pp. 374-8.

Chen, Matthew Y. (1986). "Wuxi tone sandhi: from last to first syllable dominance." Paper presented at the Conference on Languages and Dialects of China, Berkeley, California.

Clark, J. E. and C. Yallop (1990). *An Introduction to Phonetics and Phonology*, Oxford: Basil Blackwell.

Clements, Nick (1977). "Four tones from three: the extra-high tone in Anlo Ewe" in *Language and Linguistic Problems in Africa*, Ed. P.F.A. Kotey & H. Der-Houssikian. Hornbeam Press: Columbia, SC, pp. 168-191.

Clements, Nick, Alexis Michaud and Cédric Patin (2011). "Do we need tone features?" in *Tones and Features*, J. A. Goldsmith, E. Hume and L. Wetzels (eds.), Mouton de Gruyter.

Clumeck, Harold (1980). "The acquisition of tone." In *Child Phonology*, Vol. 1 (Production), Ed. G.H. Yeni-komshian, J.F. Kavanagh and C.A. Ferguson. New York: Academic Press. pp. 257-75.

DiCanio, Christian T. (2009). "The phonetics of register in Takhian Thong Chong," J. Int. Phonetic Assoc. 39, 162-188.

DiCanio, Christian T. (2012). "Coarticulation between tone and glottal consonants in Itunyoso Trique," J. Phonetics 40, 162-176.

DiCanio, Christian T. (2013). "Cue weight in the perception of Trique glottal consonants," J. Acoust. Soc. Am., Vol. 135, No. 2, February 2014.

Diffloth, Gérard (1989). "Proto-Austroasiatic Creaky Voice", *Mon-Khmer Studies* 15:139-154.

Duanmu, San (1994). "Against Contour Tone Units" *Linguistic Inquiry*, Vol. 25, No. 4 (Autumn 1994), pp.555-608. The MIT Press.

Edmondson, Jerold A. and Esling, John H. (2006). "The Valves of the Throat and Their Functioning in Tone, Vocal Register and Stress: Laryngoscopic Case Studies", *Phonology*, Vol. 23, No. 2, Between Stress and Tone (2006), pp. 157-191.

Edmondson, Jerold A. and Gregerson, Kenneth J. (1993). "Western Cham as a register language." Oceanic Linguistics Special Publication No. 24: *Tonality in Austronesian Languages*. pp.91-106.

Edmondson, Jerold A. et al. (2001). "The aryepiglottic folds and voice quality in the Yi and Bai languages: Laryngoscopic case studies" in *The Mon-Khmer Studies Journal*, vol. 31, pp. 83-100.

Egerod, Søren (1970). "Distinctive Features and Phonological Reconstructions" *Journal of the American Oriental Society*, Vol. 90, No. 1 (Jan.-Mar. 1970), pp.67-73. American Oriental Society.

Epps, Patience (2008). *A Grammar of Hup*. Mouton de Gruyter, Berlin.

Fujisaki, Hiroya, Wentao Gu and Sumio Ohno (2007). "Physiological and Physical Bases for the Command-Response Model for Generating Fundamental Frequency Contours in Tone Languages"in *Experimental Approaches to Phonology*, Ed. M-J. Solé et al. pp. 93-103. Oxford University Press.

Gage, William W. (1985). "Glottal Stops and Vietnamese Tonogenesis" *Oceanic Linguistics Special Publications*, No. 20, For Gordon H. Fairbanks, pp.21-36. University of Hawai'i Press.

Gandour, Jackson T. (1978). "The Perception of Tone" in *Tone: A Linguistic Survey*. Ed. Victoria A. Fromkin. Academic Press. pp. 41-76.

Garallek, Marc (2013). *Production and perception of glottal stops*. Diss. Ph.D. University of California, Los Angeles.

Gerfen, Chip and Kirk Baker (2005). "The production and perception of laryngealized vowels in Coatzospan Mixtec" in *Journal of Phonetics*, Vol. 33, pp. 311-334.

Gill, H. S. and H. A. Gleason (1969). *A reference grammar of Punjabi*. Patiala.

Goldsmith, John (1976). *Autosegmental Phonology*. Diss. Ph.D. MIT.

Gordon, Matthew (2001). "A typology of contour tone restrictions" *Studies in Language* 25, pp.405-444. John Benjamins Publishing Company, Amsterdam.

Greenberg, Joseph H. (1970). "Some Generalizations concerning Glottalic Consonants, Especially Implosives" in *International Journal of American Linguistics*, Vol. 36, No. 2, Hans Wolff Memorial Issue: Fascicle II (Apr., 1970), pp. 123-145.

Gussenhoven, Carlos (2000). "On the origin and development of the Central Franconian tone contrast" in *Analogy, Levelling, Markedness: Principles of Change in Phonology and Morphology*, Ed. A. Lahiri. Mouton de Gruyter, Berlin. pp. 215-260.

Gussenhoven, Carlos (2004). *The Phonology of Tone and Intonation*. Research Surveys in Linguistics. Cambridge University Press.

Gussenhoven, Carlos (2013). "From Cologne to Arzbach: An account of the Franconian 'tone reversal'" in *Nordic Prosody,* Ed. E. Asu & P. Lippus. Peter Lang: Frankfurt am Main. pp. 11-24.

Gussenhoven, Carlos and Jörg Peters (2004). "A tonal analysis of Cologne *Schärfung*" *Phonology*, Vol. 21, No. 2. Pp.251-285. Cambridge University Press.

Haraguchi, Shosuke (1977). *The Tone Pattern of Japanese: An Autosegmental Theory of Phonology*. Diss., Ph.D. MIT.

Harbaugh, Rick (1999). *Chinese Characters: A Genealogy and Dictionary (English and Mandarin Edition)*. Yale University Press. Accessed on www.zhongwen.com.

Haudricourt, André-G. (1954). "De l'origine des tons en viètnamien" *Journal Asiatique*, Vol. 242, pp.69-82. La Société Asiatique.

Hayes, Bruce (2009). *Introductory Phonology*. Blackwell.

Hillenbrand, James M. and Robert A. Houde (1996). "Role of F_0 and Amplitude in the Perception of Intervocalic Glottal Stops", *Journal of Speech and Hearing Research*, Vol. 39 (Dec. 1996), pp.1182-1190.

Hollenbach, B.E. (1977). "Phonetic vs. phonemic correspondence in two Trique dialects" in *Studies in Otomanguean Phonology*, Ed. W. Merrifield, pp. 35-68.

Hollien, H., Moore, P., Wendahl, R. W. and Michel, J. F. (1966). “On the nature of vocal fry”, *Journal of Speech and Hearing Research* 11:600-4.

Hombert, Jean-Marie (1975a). “The perception of contour tones” in *Proceedings of the 1st Annual Meeting of the Berkeley Linguistics Society* (1975). pp. 221-232.

Hombert, Jean-Marie (1975b). *Towards a theory of tonogenesis: An empirical, physiologically, and perceptually-based account of the development of tonal contrasts in language*. Diss. Ph.D. UC Berkeley.

Hombert, Jean-Marie (1976). “The Effect of Aspiration on the Fundamental Frequency of the Following Vowel” in *Proceedings of the 2nd Annual Meeting of the Berkeley Linguistics Society* (1976), pp. 212-219.

Hombert, Jean-Marie (1978a). “A model of tone systems.” in *Elements of Tone, Stress and Intonation*, Ed. D. J. Napoli.

Hombert, Jean-Marie (1978b). “Consonant Types, Vowel Quality and Tone” *Tone: A Linguistic Survey*. Ed. Victoria A. Fromkin. Academic Press. pp. 77-112.

Hombert, Jean-Marie (1984). “Tonogenesis revisited” in *Pholia*, Vol. 1, pp. 77-86.

Hombert, Jean-Marie, John J. Ohala and William G. Ewan (1979). “Phonetic Explanations for the Development of Tones,” *Language*, Vol. 55, No. 1 (Mar., 1979), pp.37-58.

Hong, Sunyoung (2007). *The phonetics of Daegu Korean lexical prosody*. Diss. Ph.D., SUNY Buffalo.

Hsieh, Feng-Fan and Michael J. Kenstowicz (2008). “Phonetic Knowledge in Tonal Adaptation: Mandarin and English Loanwords in Lhasa Tibetan” in *Journal of East Asian Linguistics*, Vol. 17, No. 4, The Loanword Phonology of East Asian Languages (Dec 2008), pp. 279-297.

Hyman, Larry M. (1978). “Historical Tonology” *Tone: A Linguistic Survey*. Ed. Victoria A. Fromkin. Academic Press.

Hyman, Larry M. (2006). “Word-Prosodic Typology” in *Phonology*, Vol. 23, No. 2, Between Stress and Tone, pp. 225-257.

Hyman, Larry M. (2009). “How (not) to do phonological typology: the case of pitch-accent” in *Language Sciences*, Vol. 31, pp. 213-238.

Hyman, Larry M. and Kobepa, Niko (2013). “On the Analysis of Tone in Mee (Ekari, Ekagi, Kapauku)” in *Oceanic Linguistics*, Vol. 52, No. 2 (Dec 2013), University of Hawai’i Press, pp. 307-317.

Hyman, Larry M. and Schuh, Russel G. (1974). “Universals of Tone Rules: Evidence from West Africa” in *Linguistic Inquiry*, Vol. 5, No. 1 (Winter 1974), pp. 81-115.

Iosad, Pavel (to appear). "'Pitch accent' and prosodic structure in Scottish Gaelic: reassessing the role of contact." in *New trends in Nordic and general linguistics*, Ed. M. Hilpert et al. Berlin: Mouton de Gruyter.

Johnsen, Sverre S. (2012). "A diachronic account of phonological unnaturalness" in *Phonology*, Vol. 29 (2012), pp. 505-531.

Kingston, John (2005). "The Phonetics of Athabaskan Tonogenesis" in *Athabaskan Prosody*, Edited by Sharon Hargus and Keren Rice. John Benjamins Publishing Company.

Klatt, Dennis H. (1973). "Discrimination of fundamental frequency contours in synthetic speech: Implications for models of pitch perception" in *Journal of the Acoustical Society of America*, Vol. 53, pp. 8-16.

Ladefoged, Peter and Ian Maddieson (1996). *The Sounds of the World's Languages*. Blackwell.

Ladefoged, Peter et al. (1998). "Phonetic structures of Scottish Gaelic," *Journal of the International Phonetic Association*. Vol. 28, pp. 1-41.

Lahiri, Aditi, Allison Wetterlin and Elisabet Jönsson-Steiner (2005). "Lexical specification of tone in North Germanic" in *Nordic Journal of Linguistics*, Vol. 28, No. 1, pp. 61-96.

Lass, Roger (1984). "Vowel System Universals and Typology: Prologue to Theory" in *Phonology Yearbook*, Vol. 1, pp. 75-111.

Laver, John (1980). *The Phonetic Description of Voice Quality*. Cambridge University Press.

Laver, John (1994). *Principles of Phonetics.* Cambridge University Press.

Lieberman, Phillip (1977). *Speech Physiology and Acoustic Phonetics*. MacMillan Publishing Co., NY.

Ma, Huiqin, Xiaoxiang Chen and Xin Liu (2012). *A Study of Tone Acquisition in a Mandarin-speaking Child under Three.* Paper presented at Speech Prosody, 6th International Conference, Shanghai.

Maddieson, Ian (2007). "Statistical Analysis of the Relationship between Syllable Structures, Segment Inventories and Tone Contrasts" in *Experimental Approaches to Phonology*, Ed. M-J. Solé et al. pp. 93-103. Oxford University Press.

Maddieson, Ian and Keng-Fong Pang (1993). "Tone in Utsat" *Oceanic Linguistics Special Publications,* No. 24, Tonality in Austronesian Languages, pp.75-89. University of Hawai'i Press.

Manaster-Ramer, Alexis (1986). "Genesis of Hopi Tones" *Journal of American Linguistics*, Vol. 52, No. 2 (Apr. 1986), pp.154-160. The University of Chicago Press.

Martinet, André (1939). "Un ou deux phonèmes?" in *Acta Linguistica*, Vol. 1, pp. 94-103.

Matisoff, James A. (1973). "Tonogenesis in Southeast Asia" in Southern California Occasional Papers in Linguistics No. 1, July 1973: Consonant Types and Tone.

Matisof, James A. (2006). “Genetic vs. Contact Relationship: Prosodic Diffusion in South-East Asian Languages” in *Areal Diffusion and Genetic Inheritance: Problems in Comparative Linguistics*. Ed. A. Aikhenvald and R. M. W. Dixon. Oxford University Press.

Matsumura, Akira (1988). *Daijirin*. Sanseido, Tokyo.

Mazaudon, Martine (1977). *Tibeto-Burman Tonogenetics. Linguistics of the Tibeto-Burman Area* Vol. 3.2.

Moreton, Elliott (2009). “Underphonologization and Modularity Bias” *Phonological Argumentation: Essays on Evidence and Motivation.* London: Equinox.

Newman, Paul (1995). “Hausa tonology: complexities in an ‘easy’ tone language” in *The Handbook of Phonological Theory*, Ed. J. Goldsmith. Basil Blackwell; Cambridge, MA.

Nhàn, Ngô Thanh (1992). *The syllabeme and patterns of word formation in Vietnamese*. NYU Ph.D. dissertation. 2/1984.

Norman, Jerry (1988). *Chinese*. Cambridge Language Surveys. Cambridge University Press.

Ohala, John J. (1978). “Production of Tone” in *Tone: A Linguistic Survey*. Ed. Victoria A. Fromkin. Academic Press. pp. 5-40.

Ozanne-Rivierre, François (1995). “Structural Changes in the Languages of Northern New Caledonia”, *Oceanic Linguistics*, Vol. 34, No. 1 (Jun., 1995), pp. 44-72.

Pierrehumbert, Janet B. and Beckman, Mary E. (1988). *Japanese Tone Structure,* Linguistic Inquiry Monograph 15, MIT Press, Cambrdige, MA.

Pierrehumbert, Janet B. and Frisch, Stefan (1997). “Synthesizing Allophonic Glottalization” in *Progress in Speech Synthesis*, Ed. J.P.H. van Santen et al. pp. 9-26. Springer-Verlag, New York.

Pike, Kenneth L. (1947). “On the Phonemic Status of English Diphthongs” in *Language*, Vol. 23 (Apr-Jun 1947), pp. 151-9.

Pulleyblank, Douglas G. (1986). *Tone in Lexical Phonology*. Springer.

Ramsey, Samuel R. (1979). “The Old Kyoto Dialect and the Historical Development of Japanese Accent.” in *Harvard Journal of Asiatic Studies*, Vol. 39, No. 1 (Jun 1979), pp. 157-175.

Riad, Tomas (1996). “Remarks on the Scandinavian tone accent typology” in *Nordlyd: Tromsø University Working Papers on Language and Linguistics*, Vol. 24. pp. 111-128.

Riad, Tomas (1998). “Towards a Scandinavian accent typology” in *Phonology and morphology of the Germanic languages*, Ed. W. Kehrein and R. Weise. Tübingen: Max Niemeyer Verlag. pp. 77-109.

Riad, Tomas (2006). “Scandinavian accent typology” in *Sprachtypologie und Universalienforschung*, Vol. 59, No. 1, pp. 36-55.

Rivierre, Jean-Claude (1993). “Tonogenesis in New Caledonia.” Oceanic Linguistics Special Publication No. 24: *Tonality in Austronesian Languages*. pp.155-174.

Ross, Malcolm D. (1993). “Tonogenesis in the North Huon Gulf Chain.” Oceanic Linguistics Special Publication No. 24: *Tonality in Austronesian Languages*. pp.133-154.

Rothenberg, Martin (2009). “Voice Onset Time vs. Articulatory Modeling for Stop Consonants” in *Logopedics Phoniatrics Vocology*, Vol. 34 (Special Issue for the Jan Gauffin Memorial Symposium), pp. 171-180.

Schuh, Russel G. (2002) “The phonology and morphology of Bade and Ngizim.” Course. Accessed on http://www.linguistics.ucla.edu/people/schuh/Bade_Ngizim/00bade_ngizim.html

Silverman, Daniel (1995). *Phasing and Recoverability*. Diss. Ph.D. UCLA.

Stevens, Kenneth N. (2000). *Acoustic Phonetics*. MIT Press.

Stevens, Peter (1960). “Spectra of Fricative Noise in Human Speech” in *Language and Speech*, Vol. 3, 32-49.

Svantesson, Jan-Olof and David House (2006). “Tone production, tone perception and Kammu tonogenesis” *Phonology*, Vol. 23, No. 2, Between Stress and Tone, pp.309-333. Cambridge University Press.

Tang, Katrina E. (2008). *The Phonology and Phonetics of Consonant-Tone Interaction*. Diss. Ph.D. UCLA.

Tesfaye, Ashenafi and Klaus Wedekind (1994). “Aspects of Omotic Tonogenesis: Shinasha” in *Journal of Ethiopian Studies*, Vol. 27, No. 2 (Dec. 1994), pp. 1-19.

Thurgood, Graham (1993). “Phan Rang Cham and Utsat: tonogenetic themes and variants.” Oceanic Linguistics Special Publication No. 24: *Tonality in Austronesian Languages*. pp.91-106.

Thurgood, Graham (1996). “Language Contact and the Directionality of Internal Drift: The Development of Tones and Registers in Chamic” in *Language*, Vol. 72, No. 1 (Mar 1996), pp. 1-31.

Thurgood, Graham (1999). “From Ancient Cham to Modern Dialects: Two Thousand Years of Language Contact and Change: With an Appendix of Chamic Reconstructions and Loanwords”, *Oceanic Linguistics Special Publications*, No. 28.

Thurgood, Graham (2008). “Vietnamese and tonogenesis: revising the model and the analysis.” *Studies in Tai and Southeast Asian Linguistics*, edited by Jimmy G. Harris, Somsonge Burusphat and James E. Harris. Bangkok: Ek Phim Thai Co. pp 263-291.

Titze, Ingo R. (1991). “The Control of Fundamental Frequency” in *Vocal Fold Physiology*, section editor C. H. Shadle, Singular Publishing Group, San Diego, pp. 129-138.

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

Tournadre, Nicolas and Sangda Dorje (2003). *Manual of Standard Tibetan: Language and Civilization*. Snow Lion Publications: Ithaca, NY.

Wang, William S-Y. (1967). “Phonological Features of Tone” *International Journal of American Linguistics*, Vol. 33, No. 2 (Apr. 1967), pp.93-105. The University of Chicago Press.

Whalen, Doug H. and Andrea G. Levitt (1995). “The universality of intrinsic F_0 of vowels” in *Journal of Phonetics* Vol. 23, pp. 349-366.

Wijk, Nicolaas van (1939a). *Phonologie, een hoofdstuk uit de structurele taalwetenschap*. Den Haag.

Wijk, Nicolaas van (1939b). “De Plaats der Tweeklanken ei, ou, ui in het Nederlandse Phonologische Systeem” in *De Nieuwe Taalgids*, Vol. 33, pp. 251-255.

Xu, Yi (1999) “F_0 peak delay: when where and why it occurs” in *International Congress of Phonetic Sciences 1999*, Ed. J. Ohala, pp. 1881-4.

Xu, Ching X. and Yi Xu (2003) “Effects of Consonant Aspiration on Mandarin Tones” in *Journal of the International Phonetic Association* Vol. 33, pp. 165-181.

Yip, Moira (1980). *The Tonal Phonology of Chinese*. Cambridge, Massachusetts: Indiana University Linguistics Club. Published 1990 by Garland Publishing, New York.

Yip, Moira (1989). “Contour Tones” *Phonology*, Vol. 6, No. 1, pp.149-174. Cambridge University Press.

Yip, Moira (2002). *Tone*. Cambridge Textbooks in Linguistics. Cambridge University Press.

Zonneveld, Wim and Mieke Trommelen (2002). “Egg, Onion, Ouch! On the Representation of Dutch Diphthongs” in *Studies in Dutch Phonology*, Vol. 4, pp. 265-292.