

## The Plausibility of Phonetic Explanations of Tonogenesis

*Arthur S. Abramson*

### 1. Background

#### 1.1 The problem

The existence of phonologically distinctive tones with pitch as their major characteristic in so many languages of the world has made many linguists, especially Western ones, wonder about the origin of tones. Perhaps linguists who speak tone languages natively wonder why so many languages are without tones! After all, according to some views on the origins of human language (e.g., Jespersen, 1922, chap. 21), vocal communication began with prosodic expressions supplemented only later by meaningful “words” that could be seen as comprising phonetic segments. Thus, one might ask whether the first languages were not all tonal. If this were indeed the situation, our discussion would not concern the origin of tones but rather the splitting of original tonal categories into a larger number of tones.<sup>1</sup> One would also be interested in the reasons for the eventual loss of tones in the ancestors of the rest of our present-day language families.

Even if, however, such early embryonic “tones” existed, we might speculate that they eventually died out during the expansion of the lexicon and gradual increase of arbitrary links between forms and their meanings, leaving the proto-languages of our present families of tone languages in a pristine state of tonelessness from which ultimately there arose tones that remained and ultimately multiplied. One way or another, it seems to be the dominant view in the literature that tones arose from a toneless state. It is this outlook that led to the widely used term *tonogenesis*, which was apparently coined by James Matisoff (1970, 1973). The concept of tonogenesis has become a broad rubric to cover not only the rise of tones in the first place but also the subsequent splitting of tones into more tones. Naturally, the historical linguist must also consider mechanisms that might underlie the subsequent merger of tones into fewer categories. The general physiological and acoustical conditions leading to tonogenesis ought to be the same wherever tonal systems have arisen, even if from area to area and language to language there are differences in detail arising, perhaps, from both internal and external linguistic factors. My emphasis here will be on languages of Southeast Asia.

#### 1.2 Tone languages

What is a tone language? Although in the realm of abstract phonology there may conceivably be some difficulty in typologically differentiating tone languages from such pitch-accent languages as Japanese and Swedish (e.g., McCawley, 1978), I accept the normal convention among linguists working with East Asian languages that the phonology of a tone language includes a number of pitch levels or contours that, along with segmental phonemes, serve to differentiate morphemes.

---

<sup>1</sup> Regardless of the ultimate origin of tones, it is certainly true that splits as well as mergers of tones have taken place in the history of languages.

Each syllable is potentially a tone-bearer, even though the lexicon may contain words with certain toneless syllables. Thus, the child in such a speech community must learn for each new word not only a string of supraglottal gestures with various excitation sources (voicing, turbulence, *etc.*) but also distinctive states and movements of pitch. That is, in such a language relative pitch is taken to be the major auditory component of tones, even though there may also be audible concomitant properties of one or more tones of the system, such as variations in loudness, syllable length, or voice quality. We must bear in mind the possibility that such concomitant factors could play a role in the emergence and development of tonal systems. Indeed, a language might hover diachronically on the borderline between a stage with distinctive phonation types or voice registers and a stage with distinctive tones.

### 1.3 Early reconstruction

It is truly impressive that using the techniques of comparative linguistics, scholars began linking the rise of tones in Asian language families with certain presumed phonetic properties of initial and final consonants in the ancestral languages. Henri Maspero (1911, 1912) did so at a time when experimental phonetic data bearing on the topic were not yet available. In spite of André-G. Haudricourt's scientific background in botany, it apparently did not occur to him to look into the existing literature on the production and perception of speech. At the time of his paper (1954) on the rise of tones in Vietnamese some research providing acoustic underpinnings to his arguments had been done (*e.g.*, House & Fairbanks, 1953), but, as far as I know, he never showed awareness of such research in his later writings.

The culmination of the research of this early period is surely the study by Haudricourt on the origin of the tones of Vietnamese. The gist of his reasoning, somewhat clarified by James Matisoff's discussion (1973, 74–76), is as follows. At the beginning of the Common Era Vietnamese was toneless. As shown in Table 1, there were three syllable types that ultimately gave rise to tones; all three had distinctive voicing in initial stop consonants: (1) open syllables, *i.e.*, those ending in a vowel or nasal, (2) syllables formerly checked with a final voiceless spirant that had become /h/, and (3) syllables formerly checked with some kind of stop, symbolized by /X/, that had become glottal stop. As for the final /h/, we might suppose that with loosening and ultimate losing of the lingual constriction of the original spirant, the somewhat open glottis of the spirant remained with a shift to glottal turbulence from the old local turbulence in the oral cavity. In addition, Haudricourt may have conjectured that the old oral stops had a simultaneous glottal closure, a characteristic claimed for a number of Southeast Asian languages (Harris, 2001). With loosening and ultimate losing of the oral closures of the stops, the glottal stop must have remained. He does not go into such detail.

Table 1: Vietnamese at the beginning of the Common Era. Adapted from Haudricourt, (1954, 80–81) with help from Matisoff (1973).

Open Syllable	Final /s/ > /h/	Final Stop (/X/) > /ʔ/
pa	pas > pah	paX > paʔ
ba	bas > bah	baX > baʔ

The emergence of three tones in Vietnamese by the end of the sixth century is apparent to Haudricourt. As shown in Table 2, characteristics of the syllable endings given in Table 1 created two contour tones as deviations from an uninflected level, the mid tone. Before disappearing, the final aspiration of the second column of Table 1 caused relaxation of the vocal folds (“le relâchement des cordes vocales”) and a concomitant lowering of the pitch of the preceding voiced portion of the vowel. On the other hand, before dropping out of the system, the final glottal closure

of the third column of Table 1 led to an increase in tension of the vocal folds (“l’augmentation de tension des cordes vocales”) with a resulting rising pitch. With the loss of the old syllable-endings the remaining pitch features took over the role of phonological distinctiveness. For this stage of the language the voicing distinction in initial position is still in place.

Table 2: Vietnamese by the end of the 6th century. Voicing distinction preserved; emergence of three tones. Adapted from Haudricourt (1954, 80–81) with help from Matisoff (1973).

Tones		
Mid	Falling	Rising
pa	pâ	pǎ
ba	bâ	bǎ

With support from Maspero’s earlier work, Haudricourt goes on to point out that by the end of the 12th century the three tones had been doubled to six because of the loss of the consonantal voicing distinction in word-initial position. That is, these linguists supposed that upon release of the old voiced stop consonants the pitch at the onset of the vowel was lower than that at the release of the old voiceless stops. It had to be assumed that such pitch perturbations were sufficiently audible to be enhanced during the time of the merging of the voicing categories. The enhancement was perhaps facilitated by the earlier emergence of contour tones from the loss of syllable-final aspiration and glottal stop. The doubling of the tones, then, served to preserve the phonological distinction between two lexical classes previously distinguished by the old voicing contrast in initial position. These phenomena, it is argued by our sources, hold not only for Vietnamese but also for the Thai and Chinese families.

#### 1.4 Indirect support from speech acoustics

As we have seen, very provocative assertions about the role of phonetic features at the beginnings and endings of syllables in tonogenesis have been made by historical linguists, who did this without any supporting physiological or acoustical data. At this point, let us consider what evidence, essentially acoustic, existed in the time of Haudricourt’s writings and not long thereafter.

Although there were precursors, the pioneering seminal work on the topic was surely the 1953 study by Arthur S. House and Grant Fairbanks. They had young adults, native speakers of American English at the University of Illinois, record many utterances of six English vowels with identical initial and final consonants, English /ptkfsbdgvzmn/. Thus, the voicing, oral–nasal, stop–fricative, and place-of-articulation contrasts were represented, with each syllable beginning and ending with the same consonant. The vocalic properties measured were duration, fundamental frequency ( $F_0$ ), and relative power. Of concern to us here is their finding that vowels in a voiceless environment had a higher  $F_0$  (109, Figure 2).<sup>2</sup> Unfortunately for us, they do not give us a clear picture of possible independent effects of initial and final factors. They point out (106) that a search of the available literature failed to show any effect of syllable-finals on  $F_0$ ; however, in a recent study of Italian (Esposito, 2002) there was no evidence of an effect of syllable-final consonantal voicing on preceding vowels.<sup>3</sup> Also in another paper (Lea, 1973), acoustic evidence of no effect on  $F_0$  of final segments is presented. Indeed, although Wayne Lea worked with American English, he succeeded in including glottal stop but not final [h] as one of the final segments. A study that does

<sup>2</sup> Also, high vowels tend to have a higher  $F_0$  than low vowels.

<sup>3</sup> Although canonical shapes of Italian words are not normally expected to end in stop consonants, Esposito assures her readers that the language has many commonly used such loan words.

include final glottal stop and [h] (Mohr, 1971), in utterances recorded by phonetically trained native speakers of Chinese, Russian, and German, no effects of any final consonants on  $F_0$  of preceding vowels were observed. In his Table IV (72) Mohr puts [h] and [ʔ] together, implying that they do not differ from each other. Hombert (1978) looked at the matter more closely in the utterances of four native speakers of Arabic and found (93–95), as posited by Haudricourt, rising  $F_0$  contours for final glottal stop and falling ones for final voiceless aspiration. Hombert also demonstrated psychoacoustically that such rising and falling  $F_0$  movements are audible.

Another noteworthy early acoustic study (Lehiste & Peterson, 1961) sought to determine what  $F_0$  perturbations should not be included in studies of sentence intonation. The results gave one more confirmation of the higher and lower  $F_0$  values upon the release of voiceless and voiced consonants respectively. Thus, in the clearest cases there is an  $F_0$  fall toward the intonation level of the utterance after a voiceless consonant and a rise after a voiced consonant. Some studies (e.g., Kohler, 1985; Ohde, 1984), however, do not support a dependable dichotomy between fall and rise.  $F_0$  upon release of a voiced stop may be on a level with, or at least not separable from, the rest of the contour.

The foregoing acoustic findings demonstrated a clear correlation between voicing states of initial consonants and the  $F_0$  heights of accompanying vowels. At least one piece of research lends credence to the posited effect of syllable-final elements. Fairly early in this program of research into tonogenesis, then, there seemed to be a substantive base for the logic of the historical linguists.

### 1.5 $F_0$ as an acoustic cue to voicing distinctions

Whatever the physiological or aerodynamic mechanisms might be that cause a correlation of  $F_0$  height with environing voicing states, the question remains as to how detectable such differences in the vibration rate of the vocal folds would have to be to allow for enhancement to the status of phonological tones. That is, the  $F_0$  perturbed upward somewhat by voicelessness and perhaps downward by voicing would have to be large enough to be heard as a pitch difference, which would ultimately take on a distinctive role as the voicing distinction itself died out. Some experiments in the literature addressed this question indirectly by testing whether such perturbations could be acoustic cues to consonantal voicing distinctions. Early perceptual studies with American English listeners (Haggard et al., 1970) and both American English and Japanese listeners (Fujimura, 1971) showed that shifts in  $F_0$  after the release of initial stops in synthetic speech could influence consonantal voicing judgments. This matter will be further addressed in the latter part of this paper, but at this point we must at least concede that the *perceptual* plausibility of at least one phonetic explanation of tonogenesis has not been undermined.

### 1.6 Representative publications on tonogenesis

In section 1.3 the groundbreaking work by three scholars, Maspero, Haudricourt, and Matisoff, was mentioned. Let us return briefly to Matisoff (1973, 77–81) for his compelling argument on monosyllabicity as a requirement for potential tonogenesis. If the laryngeal mechanisms at work here are universal, he asks, why is it that all languages have not been tonal at some time? He argues, "...it appears that to become truly tonal a language must have a basically *monosyllabic* structure (i.e., the morpheme must be only one syllable long)." He goes on to say, "There is something about the tightly structured nature of the syllable in monosyllabic languages which favors the shift in contrastive function from one phonological feature of the syllable to another." In such polysyllabic languages as Japanese and Swedish there may develop pitch-accent systems, which differ from true tonal systems. Finding Matisoff's requirement of monosyllabicity an oversimplification, Timothy Light (1978, 119) uses his data from Hakka Chinese to argue for the importance of "the loss of positions in the canonical syllable shape." He writes (123), "In this

revised version of tonogenesis, then, tones come about because of a cluster of losses combined with constraints that will not permit these losses to be compensated for on the segmental level. Metaphorically, the language has ‘nowhere to go’ but to the suprasegmental level if it is to maintain the order of magnitude of distinctions it has had.” That is, Light does not reject the importance of monosyllabicity but claims that intersecting with it is the foregoing factor.

Haudricourt points out that the kind of tonogenesis posited for Vietnamese was also present in other language families of southeast Asia, among them the Tai family, which includes two national languages, Thai of Thailand and Lao of Laos, as well as such minority languages as Shan of Myanmar (Burma). The primary source on proto-Tai (Li, 1971) admits from comparative analysis of Tai languages of the likelihood of a first stage caused by the effects of final consonants that were then lost but warns us (24), “...it is impossible to recover the final consonants that are assumed to have been dropped. If such a hypothesis is tenable, it must refer to a stage of the language prior to proto-Tai, and we are not yet ready to specify in detail the phonological system of such an early stage.”<sup>4</sup> Li, seeing no problem with the second stage of tonogenesis, goes on to say (25), “From a comparative study of the tones of the modern dialects, it is apparent that the opposition of voicing and voicelessness of the initial consonant influences practically all the tones in all dialects.”

When Paul K. Benedict began publishing his work (1942) on putative genetic relations between Thai and Austronesian through an intermediary Kadai group of languages, he had to face the common understanding that Thai was tonal but Austronesian, as exemplified by Malay and Tagalog, was not. It turns out, however, that several studies (e.g., papers in Edmondson & Gregerson, 1993) now show that tonogenesis has taken place in some Austronesian languages. In one of them Graham Thurgood (1993, 93) asserts that in Phan Rang Cham and Utsat only initial voiced obstruents yielded tone-lowering, making “it clear that it was not voicing *per se* that led to low tone, but rather there was perhaps a two-step process: first, the voiced obstruents led to breathy voice, and then the breathy voice led to low tone.”

Thurgood’s point about the role of phonation type in tonogenesis is elaborated in an important contribution by him (2002). Returning, as everyone does, to Haudricourt’s classic consonant-based model,<sup>5</sup> Thurgood develops a model based on properties of laryngeal control. One such crucial property is phonation type with voice quality as its principal auditory correlate. Two departures from modal (clear or normal) voice of concern here are creaky voice and breathy voice. In some languages these phonation types are grouped with other phonetic properties to form phonologically distinctive register complexes. These complexes may be distinguished by differences in such features as voice quality, pitch, and vowel quality. In his broad overview of both the historical linguistic and the instrumental phonetic literature Thurgood concludes that not voiced sonorants or fricatives normally but voiced obstruents, since only they are likely to impart breathy voice to following vowels, bring about lowering of pitch. He believes that his model is better able to handle tonogenesis not only in Southeast Asia but also elsewhere in the world. Supporting findings have been reported by Diffloth (1982, 1984).

Although the complicated consequences of the phenomenon of tonogenesis are inevitably subjected to rather abstract phonological analysis within one theory or another, the matter is not of central concern here; nevertheless, I should like to mention in passing one treatment (Brunelle, 1999) that posits binary laryngeal distinctive features in a framework of feature geometry. It has the merit of taking phonetic evidence very seriously. I hasten to add that such reasoning may not be pertinent to

<sup>4</sup> As far as I know, the proposed genetic link of Tai with Austronesian (Benedict, 1942) is still controversial. If, of course, the evidence becomes more convincing, we may learn more about the stage before proto-Tai.

<sup>5</sup> Thurgood (2002, 333, f.n. 1) does say that as a neo-Praguian, Haudricourt was surely not strongly wedded to a segmental view.

our understanding of the rise of tones from a pristine state of tonelessness and their multiplication in later stages. A stimulating source for a critical discussion of these phonological matters is a book by Anthony Fox (2000).

## 2. Phonetic rationale

In sections 1.4 and 1.5 I presented preliminary information on phonetic underpinnings for the general reasoning about tonogenesis. Let us return to the topic now for a more comprehensive treatment. An excellent early source for research into the topic is the survey by Hombert, Ohala, and Ewan (1979). With a focus on the widely accepted forces bearing on tonogenesis, an overview of laryngeal and aerodynamic factors, acoustics, and perception is presented. This is followed by material on other possible mechanisms not ordinarily mentioned.

### 2.1 Laryngeal and aerodynamic factors

#### 2.1.1 Syllable-final position

Most of the relevant phonetic literature grew out of concerns for matters other than tonogenesis and concentrated on the effects of syllable-initial consonantal properties on the fundamental frequency ( $F_0$ ) of following vowels. By the widely accepted view of the first stage of tonogenesis, at least in Southeast Asia (Haudricourt, 1954), before dropping out of the system Vietnamese final /h/ lowered the  $F_0$  of the preceding vowel and final /ʔ/ raised it. Supposedly this happened in other families too. Only one source that I can find (Hombert, 1978, 92–95), a substantial one at that, lends direct support to the acoustic and perceptual plausibility of this reconstruction. The few other sources on the effects of syllable-final characteristics cast doubt on the effect of final voicing distinctions but do not examine a contrast between final glottal stop and final voiceless aspiration.

Hombert does not offer a possible physiological or aerodynamic explanation of the effects. Although he does speculate on possible effects of high and low tones on following consonants (95–96), this line of thought seems to be irrelevant here, since the segment [ʔ] arose from old oral stops and [h] from [s]. We are left then with Haudricourt's suggestion that final aspiration caused relaxation of the vocal folds with resulting lowering of their vibration rate, while glottal stop caused tensing of the fold with resulting raising of the vibration rate. For the aspiration, of course, there also had to be abduction of the folds to permit the passage of turbulence through the glottis. Glottal stop requires adduction of the folds and presumably medial compression. The glottal pulses end once a tight closure is achieved. To verify such explanations would require an examination of laryngeal behavior during such utterances.

#### 2.1.2 Syllable-initial position

**Aerodynamics.** The somewhat open glottis likely to be present for voicelessness allows for a higher rate of airflow, which could induce a higher rate of glottal pulsing and thus a higher  $F_0$  at the onset of voicing. For an initial stop with voicing during its closure, however, the airflow across the glottis is intermittent in the form of glottal pulses with a consequent lowering of the rate of glottal pulsing and the resulting  $F_0$  upon release of the consonant. These effects by themselves are too short in duration to account for the full  $F_0$  perturbations that are observed (Ohala, 1973, 8–9).

**Tension of the Vocal Folds.** The larynx, our organ for producing voice while regulating the frequency and mode of vibration of the vocal folds, must play an important role in tonogenesis. Two approaches to the topic are found in the recent literature. In one of them it is the height of the larynx in the pharynx that exerts vertical tension on the vocal folds. In the other the factor is said to be essentially horizontal tension of the vocal folds.

Against a background of early reports of observations of raising of the larynx in the neck for higher

pitches of the voice and lowering for lower pitches, Ewan and Kronen (1974) used a photoelectric device, the thyroumbrometer, to track movements of the larynx in spoken English, Thai, French, and Hindi, languages that differ in the number and types of voicing categories. They did indeed find that the larynx is generally higher for voiceless stops. These data suggest that the mechanism affecting  $F_0$  must be, at least in part, vertical tension on the vocal folds (Hombert, 1978, 81). Whether or not this can be verified,<sup>6</sup> it must be recognized that lowering of the larynx seems to be part of a normal physiological strategy to enlarge the volume of the supraglottal vocal tract to help maintain a sufficient trans-glottal pressure drop to allow for the maintenance of voicing during the closure of an oral stop consonant (Bell-Berti, 1975; Nihalani, 1975; Westbury, 1983).

The argument for horizontal, *i.e.*, longitudinal, tension of the vocal folds does fit with the well-attested finding that the cricothyroid muscle controls that tension to effect changes in  $F_0$ . An early application of this factor to the matter of  $F_0$  perturbations with initial stop voicing (Halle & Stevens, 1971) proposed that “slack” vocal folds facilitate voicing and low  $F_0$ , while “stiff” vocal folds facilitate high  $F_0$  and voicelessness in consonants. These two conditions, together with “spread glottis” and “constricted glottis,” were meant to constitute a feature system. As a feature system for phonology, this model has stirred up some dissatisfaction (Anderson, 1978, 161–165). At the same time, the notion of stiffness versus slackness has merit for our question.

Strong support for longitudinal tension has come from the research of Anders Löfqvist and his collaborators (Löfqvist et al., 1989). They made electromyographic and acoustic recordings in stressed position in a carrier sentence of initial voiced and voiceless stops and fricatives recorded by a native Dutch speaker and initial voiced and voiceless stops, fricatives, and affricates recorded by two native speakers of American English. The two languages differ in that the English initial voiceless stops are aspirated in most dialects, but the Dutch stops are not.<sup>7</sup> During the occlusion or constriction of the voiceless consonants and the transition into the following vowel, a heightened contraction of the cricothyroid muscle, a tensor of the vocal folds, was found in correlation with a higher  $F_0$  of that portion of the speech signal. Aspiration made no difference, and there was very little effect on the preceding vowel. Without discounting the possible contribution of other factors in the literature, it seems clear that longitudinal tension acts as a suppressor of voicing.

## 2.2 Acoustics

### 2.2.1 Syllable-final position

We return here to the work of Hombert (1978, 92–94) for more attention to his data on the effects of final glottal stop and [h] on the  $F_0$  of the preceding vowel. Four native speakers of Arabic, a language in which the contrast is distinctive, recorded randomized nonsense syllables in a carrier sentence.  $F_0$  was measured from the offset of the vowel back every 10 msec to 100 msec before the offset.  $F_0$  was found to go up (93, Figure 7; 94, Table 2) at least 9 Hz before [ʔ] and down 25 Hz before [h]. The difference was significant at least 70 msec before the end of the vowel.

### 2.2.2 Syllable-initial position

In addition to the studies cited in section 1.4, much other acoustic research supports lower and higher onsets of  $F_0$  of the following vowel for, respectively, voiced and voiceless settings of the larynx for the initial consonant. The average time-courses of the effects are illustrated in Figure 1 adapted from Hombert (1978, 80), who had five native speakers of American English record

<sup>6</sup> As far as I know, the necessary muscular and mechanical links have not been found.

<sup>7</sup> That is, in terms of voice onset time Dutch /ptk/ have short voicing lag between the release and the onset of voicing, while English /ptk/ have long lag (Lisker & Abramson, 1964, 400).

utterances of all the English stop consonants in initial position in nonsense syllables in a carrier sentence. The curves are still significantly different 100 milliseconds after onset and then merge. His work is an important supplement to early studies that presented only averages without time-courses. Work by others (*e.g.*, Ohde, 1984) furnishes similar data for English and German (Kohler, 1985).

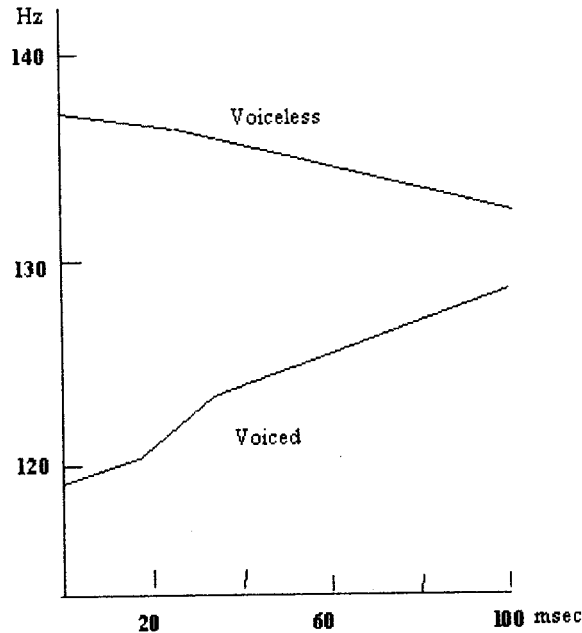


Figure 1: Average F0 contours following American English stops of five speakers. Adapted from Hombert, 1978, 80, Figure 1.

Among Southeast Asian languages, data on the effects of syllable-initial voicing states on  $F_0$  have been published for Thai (Erickson, 1975; Gandour, 1974). For another, Burmese, Maddieson (1984) exploits the presence of a voicing distinction in nasal and lateral sonorants to show that the effect of voicing states on  $F_0$  is not limited to stop consonants. This important point on the effects of phonetic properties rather than segments as such accords well with Thurgood's thesis already mentioned.

### 2.3 Perception: pitch effects from $F_0$ perturbations

The studies mentioned earlier (Haggard et al., 1970; Fujimura, 1971) demonstrated that shifts in pitch caused by  $F_0$  perturbations linked to syllable-initial voicing states could influence the perception of voicing distinctions. To this should be added the non-linguistic purely psychoacoustic experiment run by Hombert (1978, 94–95). Using his data on the effects of final [ʔ] and [h] on  $F_0$  in Arabic, he synthesized a set of versions of the vowel [i] varying only in  $F_0$  contours falling within the range of the Arabic data. The results showed that such perturbations were quite discriminable to 10 native speakers of American English. That is, the observed  $F_0$  differences are clearly above our human auditory threshold.

According to both Maspero and Haudricourt, the Tai language family is included among the Southeast Asian groups affected by the processes of tonogenesis discussed here. Two experiments conducted with synthetic speech by Abramson and Erickson (1992) were meant to address the plausibility of this claim by investigating possible interactions between phonetic properties of



initial consonants and tones in modern Thai, a member of that family. Standard Thai has five tones and, for the labial and alveolo-dental places of articulation, three voicing categories in initial stops: voiced, voiceless unaspirated, and voiceless aspirated. Since it had been shown previously (Lisker & Abramson, 1970) that these voicing categories lie acoustically and perceptually along the dimension of voice onset time (VOT), it was possible to design synthetic labial-stop stimuli that varied incrementally in VOT along with upward and downward shifts of  $F_0$  starting with the first voicing pulse after stop-release. The randomized stimuli on the mid tone with the long vowel /a:/ were played to 48 native speakers in Thailand for labeling as one of three possible Thai words. The perceptual crossover point on the VOT dimension between /b/ and /p/ judgments was significantly affected in the expected way by the  $F_0$  shifts. The boundary between /p/ and /ph/, however, was not affected. That is, the aspiration of the stimuli with delayed onsets of voicing after stop-release (voicing lag) made no difference; the main effect was on the distinction between voiced and voiceless categories (Abramson & Erickson, 1992, 9, Figure 2).

In the second experiment, meant to test the hypothesis that the voicing states of initial stop consonants should affect the boundaries on the  $F_0$  dimension between tones, 16  $F_0$  variants were pitted against three VOT values in synthetic stimuli with labial onsets followed again by the vowel /a/. The  $F_0$  variants had been shown (Abramson, 1978) to form a series that is perceptually divisible into the three static Thai tones, high, mid, and low. The VOT values were exemplars of the stops /b p ph/. The randomized stimuli were played to nine native speakers in the United States for identification. Significant interactions between the voicing states and all three tones were obtained.

### 3. Other possible mechanisms

#### 3.1 Phonation types

I have already mentioned Thurgood's arguments for phonation types or, more broadly speaking, voice registers, as a stage in tonogenesis. Detailed aerodynamic and physiological studies of phonation types used phonologically seem not to be readily accessible. Some descriptions of the ones that concern us here, breathy and creaky voice, as distinct from modal ("clear") voice, are available (e.g., Fischer-Jørgensen, 1967; Catford, 1964, 29–37; Klatt & Klatt, 1990, 621–625). These sources include the findings of others. For breathy voice the glottis is somewhat open at the rear allowing a higher flow of air but at the same time sufficient approximation of the anterior portions of the folds to yield pulses with a longer open quotient than in modal voice. The result is a kind of murmur with mixed voicing and turbulence as excitation sources. Also, perhaps because of the slacker folds needed to keep up the voicing in the presence of a glottal opening, there is likely to be a slowing down of the vibration with a consequent drop in pitch. Creaky (laryngealized) voice figures less often in discussions of tonogenesis. This involves very low frequencies of vibration of a small section of the vocal folds with a short open quotient. One can almost hear the separate pulses; it is something like drawing a stick along a picket fence.

As has been said above, phonation type in the phonology of a language is commonly accompanied by other audible features; together all of them are said to form a voice register in which the phonation type is conventionally taken to be the dominant trait. One of the physical properties of such a register is likely to be a characteristic  $F_0$  contour. From recent and current research one gets the impression that phonation type as such has become somewhat less salient in some of these languages, while pitch seems to be taking over the phonologically distinctive role. We probably have here evidence for the plausibility of phonation type as an important factor in tonogenesis.

The acoustic specifications of phonation types commonly found in voice-register languages are available in rather readable form (e.g., Klatt & Klatt, 1990; Ladefoged et al., 1988; Gordon & Ladefoged, 2001; Wayland & Jongman, 2003). In brief, there are two ways of looking at the

acoustic manifestations: (1) spectral tilt and (2) amount of noise relative to the harmonic content. Tilt, determined by the shape of the glottal pulse, is the ratio of the amplitude of the first harmonic to that of the second harmonic or the prominent harmonic of the first or a higher formant. For breathy voice, the spectrum is weaker than for modal voice in the higher frequencies; for creaky voice, it is stronger. Spectral tilt can best be determined for comparison of vowels with essentially the same formant frequencies. This condition is not met when vowel quality itself has become a feature of voice-register differentiation. For breathy voice, the noise component, caused by an opening in the glottis, may be intense enough to be quite audible. There are ways of calculating the ratio of harmonics to aspiration noise in the spectrum.

Let us consider a few present-day studies of Southeast Asian voice-register languages of the Mon-Khmer family that may be undergoing tonogenesis. Although other factors continue to separate the modal and breathy registers of Nyah Kur and Kui (Suai),  $F_0$  differences are prominent (Thongkum, 1988). In her work on Mon (1990) Thongkum finds even stronger evidence of a voice-register language on the brink of becoming tonal. An interesting array of stages of tonogenesis is to be seen in the dialects of Khmu (Premsrirat, 2001). In a project now near completion (Abramson, Thongkum and Nye, in preparation) the traditional modal and breathy voice registers of the Kuai dialect of Suai are apparently giving way to pitch. A fairly unstable contrast in spectral tilt is being replaced by a more robust difference in  $F_0$  contours, even though synthesis experiments reveal some remaining sensitivity to phonation type.

### 3.2 Initial consonant length

The gradual conversion of a word-initial distinction in consonant length to one of accent or tone must be very rare, since, in the first place, a consonantal length distinction in that position is rare. A series of studies of Pattani Malay ("Jawi"), the variety of Malay spoken in the southernmost provinces of Thailand, has culminated in a paper (Abramson, 2003) that shows  $F_0$  to be a very powerful cue to the perceptual identification of such words when the percept of duration of consonantal closure or constriction is not available, as in utterance-initial voiceless plosives. In addition, there are informal observations by field linguists of the instability of the length contrast in the speech of young villagers. Together with Christopher Court (personal communication) I tend to hear accentual salience on the first syllable of the typically disyllabic words if they begin with a long consonant.

## 4. Conclusion

The phonetic evidence reviewed here lends credence to the broad lines of the historical linguistic arguments for tonogenesis. The seeds of change are small perturbations of fundamental frequency that are not under the control of the speaker. That is, in times of relative stability these minor pitch shifts are presumably not noticed as such. They are automatic in the context of certain syllable-final or syllable-initial consonantal properties. For example, they are neither produced nor heard as part of the stress pattern of a word or as part of an intonation pattern. I do not mean that the perturbations at this stage are not audible; I simply mean that they are not under the voluntary control of the speaker and so have no independent linguistic function. At the same time, the phonologist's eager quest for distinctive features notwithstanding, it seems to be typical in the experimental phonetic literature that our ability to hear minimal phonemic distinctions seldom depends entirely on a single phonetic property. Rather, we are likely to be influenced by several cues, even if experiments show that they differ in relative power (Abramson & Lisker, 1985). Thus, if one of those concomitant properties is somehow subordinate to others, it may nevertheless be available to become stronger or even dominant during a transitional period of the weakening of the phonetic base of the old distinction. Thus, in the "tone-prone" language families of concern to us here, the weakening of the consonantal voicing contrast, probably with an intermediate stage of a

phonation-type contrast, could well have facilitated the enhancement of previously redundant pitch to tonal status.

A way of understanding this process is through the concept of parsing (Fowler & Brown, 1997). In the publications on the role of  $F_0$  as a cue to voicing distinctions, the listener parses out the  $F_0$  perturbations from the intonational, accentual, or tonal contour of the linguistic expression and uses them only for aid in identifying the voicing state of the consonant. Such a perturbation is ignored, it would seem, in responding to the overall pitch contour of the utterance. It, however, is ripe for advancement to a new phonological function if the language, for one reason or another, begins to lose the consonantal distinction with which the perturbation has been linked and the language is threatened with burdensome homophony.

The conviction, as expressed here, that the arguments for tonogenesis are phonetically plausible does not mean that all details have been handled. As one moves from tone language to tone language, there are puzzling instances of paradoxical changes that remain to be explained. In some cases this may reflect the skimpiness of the reconstruction, perhaps because of the paucity of historical data. As emphasized by Graham Thurgood, the adding of phonation types to our thinking has shed much light on the process. There may be other such factors waiting to be uncovered.

## 5. Acknowledgement

This research was supported by NIH Grant DC-02717 (Links between Production and Perception in Speech) to Haskins Laboratories. I am grateful to Theraphan Luangthongkum and Leigh Lisker for stimulating discussions on the topic.

## References

- Abramson, A.S. (1978). Static and Dynamic Acoustic Cues in Distinctive Tones. *Language and Speech*, 21, 319-25.
- Abramson, A.S. (2003). Acoustic Cues to Word-Initial Stop Length in Pattani Malay. *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona, Spain, 387-90.
- Abramson, A.S. & Erickson, D.M. (1992). Tone Splits and Voicing Shifts in Thai: Phonetic Plausibility. *Pan-Asiatic Linguistics: Proceedings of the Third International Symposium on Languages and Linguistics*, Vol. I, Bangkok: Chulalongkorn U, 1-16.
- Abramson, A.S. & Lisker L. (1985). Relative Power of Cues: Fo Shift Versus Voice Timing. In Victoria A. Fromkin (ed.). *Phonetic Linguistics: Essays in Honor of Peter Ladefoged*. New York: Academic Press, 25-33.
- Abramson, A.S., Thongkum, T.L., & Nye, P.W. The Acoustics of the Production and Perception of Voice Register in Suai. (In preparation).
- Anderson, S.R. (1978). Tone features. In Victoria A. Fromkin (ed.). *Tone: A Linguistic Survey*. New York: Academic Press, 133-75.
- Bell-Berti, F. (1975). Control of Pharyngeal Cavity Size for English Voiced and Voiceless Stops. *Journal of the Acoustical Society of America*, 57, 456-61.
- Benedict, P.K. (1942). Thai, Kadai, and Indonesian: A New Alignment in Southeastern Asia. *American Anthropologist*, 44, 576-601.
- Brunelle, M. (1999). Tonogénèse, registrogénèse et traits laryngaux binaires. *Cahiers de linguistique—Asie Orientale*, 28, 23-52.
- Catford, J.C. (1964). Phonation types: The Classification of Some Laryngeal Components of Speech Production. In D. Abercrombie, D.B. Fry, P.A.D., MacCarthy, N.C. Scott, and J.L.M. Trim (eds). *In Honour of Daniel Jones*. London: Longmans, 26-37.
- Diffloth, G. (1982). Registres, dévoisement, timbres vocaliques: Leur histoire en katouique. *Mon-Khmer Studies*, 11, 47-82.

- Diffloth, G. (1984). *The Dvaravati Old Mon Language and Nyah Kur*. Monic Studies. 1. Bangkok: Chulalongkorn U Printing House.
- Edmondson, J.A. & Gregerson K.J., (eds.). (1993). *Tonality in Austronesian Languages*. Oceanic Linguistics Special Publication 24. Honolulu: U of Hawaii P.
- Erickson, D. (1975). Phonetic Implications for an Historical Account of Tonogenesis in Thai. In J. G. Harris, and J. R. Chamberlain (eds.). *Studies in Tai Linguistics in Honor of William J. Gedney*. Bangkok: Central Institute of English Language, Office of State Universities, 100–11.
- Espósito, A. (2002). On Vowel Height and Consonantal Voicing Effects: Data from Italian. *Phonetica*, 59, 197–231.
- Ewan, W.G. & Kronen R. (1974). Measuring Larynx Movement Using the Thyroumbrometer. *Journal of Phonetics*, 2, 327–35.
- Fischer-Jørgensen, E. (1967). Phonetic Analysis of Breathy (Murmured) Vowels in Gujarati. *Indian Linguistics*, 28, 71–139.
- Fowler, C.A. & Brown J.M. (1997). Intrinsic  $F_0$  Differences in Spoken and Sung Vowels and Their Perception by Listeners. *Perception & Psychophysics*, 59, 729–38.
- Fox, A. (2000). *Prosodic Features and Prosodic Structure*. Oxford: Oxford U Press.
- Fujimura, O. (1971). Remarks on Stop Consonants: Synthesis Experiments and Acoustic Cues. In L.L. Hammerich, Roman Jakobson, & E. Zwirner (eds.). *Form and Substance: Phonetic and Linguistic Papers Presented to Eli Fischer-Jørgensen*. Copenhagen: Akademisk, 221–32.
- Gandour, J. (1974). Consonant Types and Tone in Siamese. *Journal of Phonetics*, 2, 337–50.
- Gordon, M. & Ladefoged P. (2001). Phonation Types: A Cross-Language Overview. *Journal of Phonetics*, 29, 383–406.
- Haggard, M., Ambler S. & Callow M.(1970). Pitch as a Voicing Cue. *Journal of the Acoustical Society of America*, 47, 613–17.
- Halle, M. & Stevens K.N. (1971). A Note on Laryngeal Features. *Quarterly Progress Report of the Research Laboratory of Electronics MIT*, 101, 198–213.
- Harris, J.G. (2001). States of the Glottis of Thai Voiceless Stops and Affricates. In M.R. Kalaya Tingsabadh, and Arthur S. Abramson. (eds.). *Essays in Tai Linguistics*. Bangkok: Chulalongkorn University Press, 3–11.
- Haudricourt, A.G. (1954). De l'origine des tons en vietnamien. *Journal asiatique*, 242, 69–82.
- Hombert, J.-M. (1978). Consonant types, vowel quality, and tone. In Victoria A. Fromkin (ed.). *Tone: A Linguistic Survey*. New York: Academic Press, 77–109.
- Hombert, J.-M., Ohala J. J. & Ewan W.G. (1979). Phonetic Explanations for the Development of Tones. *Language*, 55, 37–58.
- House, A.S. & Fairbanks G. (1953). The Influence of Consonant Environment upon the Secondary Acoustical Characteristics of Vowels. *Journal of the Acoustical Society of America*, 25, 105–13.
- Jespersen, O (1922). *Language: Its Nature, Development and Origin*. London: George Allen & Unwin.
- Klatt, D.H. & Klatt. L.C. (1990). Analysis, Synthesis, and Perception of Voice Quality Variations among Female and Male Speakers. *Journal of the Acoustical Society of America*, 87, 820–57.
- Kohler, K.J. (1985).  $F_0$  in the production of lenis and fortis plosives. *Phonetica*, 39, 199–218.
- Ladefoged, P., Maddieson I. & Jackson M. (1988). Investigating Phonation Types in Different Languages. In Osamu Fujimura (ed). *Vocal Physiology: Voice Production, Mechanisms and Functions*. New York: Raven Press, 297–317.
- Lea, W.A. (1973). Segmental and Suprasegmental Influences on Fundamental Frequency Contours. In Larry M. Hyman (ed). *Consonant Types & Tone*. Southern California Occasional Papers in Linguistics, No. 1. Los Angeles: Linguistics Program, U of Southern California, 16–70.
- Lehiste, I. & Peterson G. (1961). Some Basic Considerations in the Analysis of Intonation. *Journal of the Acoustical Society of America*, 33, 419–25.

- Li, Fang Kuei. (1977). *A Handbook of Comparative Tai*. Oceanic Linguistics Special Publication No. 15. Honolulu: The University Press of Hawaii.
- Light, T. (1978). Tonogenesis: Analysis and Implications. *Lingua*, 46, 115–131.
- Lisker, L. & Abramson A.S. (1964). A Cross-Language Study of Voicing in Initial Stops: Acoustical Measurements. *Word*, 20, 384–422.
- Lisker, L. & Abramson A.S. (1970). The Voicing Dimension: Some Experiments in Comparative Phonetics. *Proceedings of the 6th International Congress of Phonetic Sciences*, Prague, 1967. Prague: Academia, 563–567. (Reprinted in *Acoustic Phonetics: A Course of Basic Reading*. D.B. Fry.(ed.). Cambridge: Cambridge University Press, 1976, 348–52.)
- Löfqvist, A., Baer, T., McGarr, N. & Story R. S. (1989). The Cricothyroid Muscle in Voicing Control. *Journal of the Acoustical Society of America*, 85, 1314–21.
- Maddieson, I. (1984). The Effects on  $F_0$  of a Voicing Distinction in Sonorants and their Implication for a Theory of Tonogenesis. *Journal of Phonetics*, 12, 9–15.
- Maspero, H. (1911). Contribution à l'étude du système phonétique des langues thai. *Bulletin de l'École française d'Extrême-Orient*, 11, 153–169.
- Maspero, H. (1912). Étude sur la phonétique historique de la langue annamite: Les initiales. *Bulletin de l'École française d'Extrême-Orient*, 12, 1–127.
- Matisoff, J.A. (1970). Glottal Dissimilation and the Lahu High-Rising Tone: A Tonogenetic Case Study. *Journal of the American Oriental Society*, 90, 13–44.
- Matisoff, J.A. (1973). Tonogenesis in Southeast Asia. In Larry M. Hyman (ed.). *Consonant Types & Tone*. Southern California Occasional Papers in Linguistics, No. 1. Los Angeles: Linguistics Program, U of Southern California, 71–95.
- McCawley, J.D. (1978). What is a Tone Language? In Victoria A. Fromkin (ed.). *Tone: A Linguistic Survey*. New York: Academic Press, 113–31.
- Mohr, B. (1971). Intrinsic Variations in the Speech Signal. *Phonetica*, 23, 65–93.
- Nihalani, P. (1975). Velopharyngeal Opening in the Formation of Voiced Stops in Sindhi. *Phonetica*, 32, 89–102.
- Ohala, J.J. (1973). The Physiology of Tone. In Larry M. Hyman (ed.). *Consonant Types & Tone*. Southern California Occasional Papers in Linguistics, No. 1. Los Angeles: Linguistics Program, U of Southern California, 1–14.
- Ohde, R. (1984). Fundamental Frequency as an Acoustic Correlate of Stop Consonant Voicing. *Journal of the Acoustical Society of America*, 75, 224–30.
- Premrsirat, S. (2001). Tonogenesis in Khmu Dialects of SEA. In Graham Thurgood (ed.). *Papers from the Ninth Annual Meeting of the Southeast Asian Linguistics Society, 1999*. Tempe, Arizona. Arizona State University Program in Southeast Asian Studies, 121–34.
- Thongkum, T.L. (1988). Phonation Types in Mon-Khmer Languages. In Osamu Fujimura (ed.). *Vocal Physiology: Voice Production, Mechanisms and Functions*. New York: Raven Press, 319–33.
- Thongkum, T.L. (1990). The Interaction between Pitch and Phonation Type in Mon: Implications for a Theory of Tonogenesis. *Mon-Khmer Studies*, 16–17, 11–24.
- Thurgood, G. (1993). Phan Rang Cham and Utsat: Tonogenetic Themes and Variants. In Jerold A Edmondson and Kenneth J. Gregerson (eds.). *Tonality in Austronesian Languages*. Oceanic Linguistics Special Publication 24. Honolulu: U of Hawaii P, 91–106.
- Thurgood, G. (2002). Vietnamese and Tonogenesis: Revising the Model and the Analysis. *Diachronica*, 19, 333–363.
- Wayland, R. & Jongman A. (2003). Acoustic Correlates of Breathy and Clear Vowels: The Case of Khmer. *Journal of Phonetics*, 31, 181–201.
- Westbury, J. (1983). Enlargement of the Supraglottal Cavity and its Relation to Stop Consonant Voicing. *Journal of the Acoustical Society of America*, 73, 1322–26.