Diachronic Tone Splits and Voicing Shifts in Thai: Some Perceptual Data

Arthur S. Abramson† and Donna M. Erickson

ABSTRACT

Proto-Tai is said to have had three phonemic tones and four consonantal voicing categories, which would have been inherited by Old Thai (Siamese). Correlations between tones and initial consonants across the Tai languages have led to the positing of tonal splits conditioned by the voicing states of initial consonants with a subsequent shifting of voicing features in certain lexical classes. Thus for each tone of Old Thai, words with initial voiced consonants developed a lower tone and words with initial voiceless consonants, a higher tone. Two types of experiment were designed to test the phonetic plausibility of the argument: (1) CV syllables were synthesized with three values of voice onset time (VOT) acceptable as Thai /b p ph/. Each of these was combined with a continuum of F0 contours that had previously been divided perceptually into the high, mid and low tones. These syllables were played to native speakers of Thai for tonal identification. (2) Labial stops with nine values of VOT separable into /b p ph/ categories were coupled on synthetic mid-tone and low-tone CV syllables with upward and downward F0 onsets varying in extent and duration. The resulting syllables were played to native speakers for identification of the initial consonants. The historical argument receives some support from the experimental data.

The term tonogenesis, apparently first used by James Matisoff (1970), can mean the emergence of phonologically distinctive tones in a previously toneless language under the influence of certain contextual features. Although in a given case the best historical reconstruction may lead to just that conclusion, we see no reason to believe that tonal distinctions should not be just as primitive as vocalic or consonantal distinctions in a protolanguage. A further use of the term tonogenesis has been as a label for the splitting of old tonal categories into a larger number of tones. A consensus of historical linguists is that such has been the development in the Tai family of languages. J. Marvin Brown (1975) refers to the "great tone split ... that swept through China and northern Southeast Asia nearly a thousand years ago."

*This paper was presented before the Linguistic Society of America, Honolulu, August 1977.

†Also the University of Connecticut.

During the period of the emergence of its daughter languages, Proto-Tai is generally said to have had three phonemic tones on "smooth" syllables (those ending in vowels, glides or nasals) and four voicing categories for initial consonants, which would all have been inherited by Thai (Siamese). With some help from the ancient writing systems, examination of correlations between tones and initial consonants has led to the positioing of tonal splits conditioned by the voicing states of initial consonants with a subsequent phonological shifting of voicing features in certain lexical classes [Maspero, 1911; Li, 1947, 1977; Haudricourt, 1956; (Gedney, see footnote 1). This development purportedly underlies the system of five tones and three consonantal voicing categories of modern Thai. Thus, ignoring the special problems of one of the four classes of consonants, the so-called glottalized consonants, we find that for each tonal category of Old Thai, words with initial voiced consonants developed a lower tone, and words with initial voiceless consonants, a higher one. Thus the three original tones would have split into six. In fact, given the vicissitudes of language change spread over related languages, we find that Central Thai, which is the dialect of the Bangkok region and the official language of Thailand, has only five tones, while other dialects and other Tai languages have six or more, with differences among them in tonal shapes as well.

Independently of these historical hypotheses, it has been known for some time that the fundamental frequency of a syllable beginning with a voiced consonant is likely to be lower than that of a syllable beginning with a voiceless consonant (House and Fairbanks, 1953; Leibiste and Peterson, 1961). We know that for Thai (Gandour, 1974; Erickson, 1975) and other languages (Hombert, 1975), voiced initials are in fact accompanied by an upward movement of fundamental frequency, and voiceless consonants, by a downward movement; both of these perturbations then tend to move back toward the prosodic norm of the syllable as a whole. While the physiological mechanisms underlying these perturbations are still rather controversial, fundamental-frequency movements of comparable magnitude have been shown by Hombert to be quite perceptible. It has also been found that either in exaggerated form (Haggard, Ambler and Callow, 1970) or within more or less normal ranges (Fujiwura, 1971; Abramson, 1974), such perturbations can influence phonemic judgments of voicing. If we assume these findings in production and perception to be universal and thus to apply to Old Thai, we might suppose that speakers of the language, already accustomed to a three-way tonal contrast, were psychologically receptive to the pitch fluctuations normally occurring with voicing distinctions. However, attention was gradually shifted to syllable nuclei as pitch perturbations were more and more enhanced in perception and production, and away from syllable initials, to the detriment of the latter. That is, the phonemicization of the pitch fluctuations, yielding an increase in tonal categories, helped to keep the old lexical classes apart even while the consonantal voicing categories, to some extent under the influence of pitch, decayed, shifted and even coalesced.

---

1Gedney, W. J. Future directions in comparative Tai linguistics. (unpublished manuscript).
Our aim was to examine the perceptual plausibility of the foregoing historical arguments. We approached the problem in two ways. First, we tested for the effects of systematic pitch perturbations on the identification of initial stop consonants. Then we experimented with the effects of the voicing states of initial stop consonants on tone identification. In order to have incremental control over the phonetic dimensions of interest to us, we followed the common practice in experimental phonetics of using synthetic speech.

Techniques of acoustic analysis and synthesis have shown that the voiced, voiceless unaspirated, and voiceless aspirated stops of Modern Thai lie along a dimension of voice onset time (VOT), namely, the temporal relation between the closing of the glottis for audible pulsing and the release of the occlusion of the initial stop (Lisker and Abramson, 1964; Abramson and Lisker, 1965). VOT itself is simply an instance in utterance-initial position of a more general phenomenon of laryngeal timing in consonant distinctions (Abramson, 1977). In Figure 1 are plotted the responses of 48 native speakers of Thai to a synthetic continuum of variants of VOT in labial stop consonants with the vowel /aa/ on the mid tone. The abscissa shows values of VOT, with voicing lead in negative numbers and voicing lag in positive numbers, while zero means voice onset at the moment of release. For voicing lead, low-frequency harmonics are present before the release during the simulated occlusion; for voicing lag, until the moment of pulsing onset after the release, the upper formants are filled with noise to simulate aspiration, and the first formant is omitted. The ordinate shows percent identification of each of the stops. As can be seen, the three expected voicing categories emerge.

Having demonstrated the sufficiency of VOT as a cue to the three-way voicing distinction in Thai, we turned to the question of the effect of pitch perturbations at syllable beginnings on consonant identification. The stimuli were made for this experiment by varying VOT and extent and duration of initial fundamental-frequency shifts. The basic syllable pattern for all stimuli was a set of formant transitions appropriate to the labial place of articulation and steady-state formants for a vowel of the type [a:]. With the data from Figure 1 as a baseline, we chose nine VOT values ranging from -100 to +80 msec to span the three voicing categories. An acceptable mid tone was produced by using a level fundamental frequency at 120 Hz with a small drop at the end. As shown in Figure 2, four shifts of fundamental frequency were applied to the beginning of the contour, with falls from 10 Hz and 20 Hz above and rises from 10 Hz and 20 Hz below. These values were derived from production data published by Erickson (1974). To these four was added a fifth variant with a level onset, that is, no shift. Finally, because of some uncertainty in the literature as to the appropriate duration of such shifts, we synthesized them with three time spans: 50, 100, and 150 msec. The resulting 117 stimuli were presented in a number of randomizations for identification as to voicing category. The results for 46 speakers of Thai for the 100-msec condition are presented in Figure 3. From top to bottom the three graphs show identifications of the stimuli as /b/, /p/, and /ph/, respectively. As shown by the coded lines, it is the boundary between the voiced and the voiceless unaspirated stops along the VOT dimension that is much affected by fundamental-frequency perturbations, while the boundary between the voiceless unaspirated stop and the voiceless aspirated stop hardly
Figure 1: Thai identifications of synthetic labial stops varying in voice onset time (VOT).

Figure 2: Initial F0 shifts for stop-voicing identification on the mid tone.
Figure 3: Effects of initial $F_0$ shifts on the identification of voicing in Thai initial stops.
Figure 4: Initial $F_0$ shifts for stop-voicing identifications on the low tone.
Figure 5: A set of $F_0$ contours yielding identifications as high, mid and low tones.

Figure 6: Identification functions for the contours of Figure 5.
Figure 7: Effects of the voicing states of initial consonants on the tonal identification of the contours in Figure 5.
varies. An analysis of variance shows these effects to be highly significant. The overall differences between the three time spans were not significant. It is for this reason that only one duration is given in Figure 3.

Thus it is clear that at least one voicing boundary can be pushed about by perturbations of fundamental frequency, but it is important to know that this phenomenon is not restricted to the mid tone used up to this point. Reasoning that three reconstructed tones of Proto-Tai might well have been phonetic approximations to the relatively static high, mid, and low tones of Central Thai, we chose the low tone for the next experiment. For this experiment a fundamental-frequency contour that had been shown to be acceptable for the low tone (Abramson, 1962, 1975) was imposed on the same syllable type; it began at 116 Hz and dropped to 96 Hz where it leveled off until the end of the syllable. Except for a limit on the duration of each perturbation to 100 msec, the stimuli for this experiment were analogous to those made on the mid tone. That is, as shown in Figure 4, there were downward shifts of fundamental frequency from 10 and 20 Hz above the starting point of the basic tonal contour and upward shifts from 10 and 20 Hz below, in addition to one unperturbed contour. The responses of eight subjects to randomizations of these stimuli are essentially the same as those in the experiment on the mid tone. The great discrepancy between numbers of subjects available for the two experiments makes it very difficult to do a detailed statistical comparison across the two tonal conditions; nevertheless, the main effects are repeated; it is only the boundary between the voiced stop and the voiceless aspirate that is affected. An extension of this research will be to try the same experiment with the high tone. Will it show the same pattern of responses? Since, in other studies the mid and low tones have been shown to be confusable under certain conditions (Abramson, 1976a), but never the high tone with either of them, it is conceivable that the high tone provides a suitable context for accompanying downward perturbations of its onset to have more of an effect on stop identification.

The foregoing data raise a question: Why is the boundary between the two voiceless categories not affected? A reasonable explanation may be that once one reached far enough into the voicing-lag part of the VOT dimension, the resulting stimuli are psychoacoustically very different from variants with lower VOT values—as very audible noise excitation of the formants is present in the signal. We think that this aspiration noise is so powerful a cue that any accompanying pitch shifts, even if audible, cannot affect labeling responses. The lack of effect of pitch perturbation on the voiceless aspirate in our data accords well with the historical observation that the voiceless aspirate of the protolanguage has persisted into Modern Thai. The shift of the voiced stop of the protolanguage to modern voiceless aspirate, however, seems to require a more indirect explanation. In this connection, it is important to note that in most other Tai languages this phoneme became a voiceless aspirate.

Thus we see from the preceding two experiments that the boundary between voiced stops and voiceless inaspirates is affected by initial fundamental-frequency perturbations. To continue our research into the question of interactions between voicing states and pitch in the history of Thai, we turned to an experiment on the effects of initial stop consonants on the identification of tones. As shown in Figure 5, we used a fan-shaped continuum
of fundamental-frequency variants with a common origin that had been shown previously (Abramson, 1976b) to be perceptually divisible into the three static tones, high, mid, and low. The tonal variants all started at 120 Hz and moved to the end points ranging from 152 to 92 Hz in four-Hz steps. The labeling responses of 31 Thai subjects to this continuum in that study are shown in Figure 6. Here we can see that they essentially divided the continuum into three tones. We synthesized syllables with VOT values appropriate to /baa paa phaa/ with all 16 tonal variants. Randomized lists of the stimuli were presented to Thai subjects for identification of both the stops and tones.

Figure 7 shows the effects of the initial stop consonants on tone identification. The graphs from top to bottom give the results for the identification of the low, mid, and high tones, respectively, as they are affected by the three voicing categories of stops. The three stops are indicated by the coded lines. Essentially, the data show that the tone identification is affected by the consonant categories, but in a somewhat paradoxical way. For the low-tone identification function, the voiced stop entails a significantly higher number of low-tone judgments than does the voiceless aspirated stop. For the high-tone identification function, the voiced stop entails a significantly higher number of high-tone judgments than do both the voiceless stops.

The paradox is the direction of the boundary shift for the voiced stop. It is at a higher fundamental frequency for the boundary between the mid and low tones, and at a lower fundamental frequency for the boundary between the mid and high tones. The best interpretation we have to offer at this time is the following. In the case of the low tone, we would expect that since the Thai speaker associates lower pitch with voiced initial consonants, he does not need as low a value of fundamental frequency to hear a low tone on the syllable beginning with /b/. In the case of the high tone, our reasoning is somewhat the opposite. That is, with the high tone, again the listener associates an inherent lowness with the /b/ consonant, but in this case, he compensates for the expected lowness by allowing syllables beginning with /b/ to be heard as high tones at a lower frequency range. This kind of interpretation assumes a difference between low and high tones in perceptual processing yet to be understood.

To summarize this experiment, we can say that the tonal boundaries are affected by the stop categories, although we do not yet understand the exact reason for the nature of this interaction. A more complicated experiment planned for the future is to combine the two approaches used here, namely to perturb the onsets of the 16 fundamental frequency variants in association with the voicing characteristics of the initial stops.

In conclusion, then, our experiments show that fundamental-frequency perturbations affect consonant categories and that consonant categories affect tone labeling of fundamental-frequency continua. Thus, our data lead us to the conclusion that, by and large, the historical arguments concerning interactions between tone splits and voicing shifts are perceptually plausible. As pitch perturbations loomed larger in the consciousness of the speakers and gradually took on phonemic status, one might expect that the voicing states of initial consonants would have been reassessed perceptually.
and rearticulated to furnish new production norms, thus helping to bring about
shifts in tone and consonant categories.

REFERENCES

Measurements and Experiments. (Bloomington: Indiana University Research
Center in Anthropology, Folklore and Linguistics, Pub. 20) [International
Journal of American Linguistics 28 (2, pt. 3)].

Abramson, A. S. (1974) Pitch in the perception of voicing states in Thai:
Diachronic implications. Linguistics Society of Meeting Handbook, Forty-
Ninth Annual Meeting. (A), I. [Also in Haskins Laboratories Status
Report on Speech Research SR-41, 165-174].

Abramson, A. S. (1975) The tones of Central Thai: Some perceptual experi-
ments. In Studies in Tai Linguistics in Honor of William J. Gedney,
ed. by J. G. Harris and J. R. Chamberlain. (Bangkok: Central Institute
of Language), 1-16.

Abramson, A. S. (1976a) Thai tones as a reference system. In Tai Linguistics
in Honor of Fang-Kuei Li, ed. by T. W. Gething, J. G. Harris and
P. Kullavanijaya. (Bangkok: Chulalongkorn University Press), 1-12.

Abramson, A. S. (1976b) Static and dynamic acoustic cues in distinctive tones.
Journal of the Acoustical Society of America 59, 842(A). [Also in

Abramson, A. S. (1977) Laryngeal timing in consonant distinctions. Phone
tica 34, 295-303.

Abramson, A. S. and L. Lisker. (1965) Voice onset time in stop consonants:
Acoustic analysis and synthesis. In Proceedings of the Fifth
International Congress of Acoustics, ed. by D. E. Commins, A51 (Lige:
G. Thone).

Brown, J. M. (1975) The great tone split: Did it work in two opposite ways?
In Studies in Tai Linguistics in Honor of William J. Gedney, ed. by
J. G. Harris and J. R. Chamberlain. (Bangkok: Central Institute
of Language), 33-48.

Erickson, Donna. (1974) Fundamental frequency contours of the tones of Stan-

Erickson, Donna. (1975) Phonetic implications for an historical account of
tonogenesis in Thai. In Studies in Tai Linguistics in Honor of William
J. Gedney, ed. by J. G. Harris and J. R. Chamberlain. (Bangkok: Central
Institute of Language), 100-111.

Fujimura, O. (1971) Remarks on stop consonants: Synthesis experiments and
acoustic cues. In Form and Substance: Phonetic and Linguistic Papers
Presented to Eli Fischer-Jürgensen, ed. by L. L. Hammerich, R. Jakobson
and E. Zwirner. (Copenhagen: Akademisk Forlag), 221-232.

2, 337-350.

of the Acoustical Society of America 47, 613-617.

Haudricourt, A. C. (1956) De la restitution des initiales dans les langues
monosyllabiques: le probleme du Thai commun. Bulletin de la Societe de
Linguistique de Paris 52, 307-322.

Hombert, J. M. (1975) Towards a theory of tonogenesis: An empirical, physio-
logically and perceptually-based account of the development of tonal
contrasts in languages. Doctoral dissertation. (Berkeley: University
of California).