# Can a perceptual experiment reflect tonogenesis in Tai?

**RUNGPAT Roengpitya***
Department of Linguistics
University of California at Berkeley

A reconstruction of the tonal systems in languages within the Tai language family (Li 1977) showed that tonogenesis (termed by Matisoff 1973) in Tai languages arose from a neutralisation of a voicing contrast in prevocalic consonants: new higher tonal variants after what had been voiceless consonants and new lower tonal variants after what had been voiced consonants. The two groups are pronounced as voiceless unaspirated stops in Lungchow (a language representing the Central group of Tai) and Po-ai (a language representing the Northern group of Tai), and as voiceless aspirated stops in Siamese or Standard Thai (a language representing the South-western group of Tai). This is illustrated below:

Proto-Tai: a voiceless aspirated dental stop **th-* (from Li 1977)

| | Lungchow | Po-ai | Siamese |
|---|---|---|---|
| 'to pull out' | - | toon | thɔɔn |

Proto-Tai: a voiceless unaspirated dental stop **t-* (from Li 1977)

| | Lungchow | Po-ai | Siamese |
|---|---|---|---|
| 'to castrate' | toon | toon | tɔɔn |

Proto-Tai: a voiced dental stop **d-* (from Li, 1977)

| | Lungchow | Po-ai | Siamese |
|---|---|---|---|
| 'section' | toon | toon | thɔɔn |

The cue to differentiate the two sets, namely, the Proto-Tai voiceless aspirated stops and the Proto-Tai voiced stops, is no longer the voicing but the fundamental frequency (Fo) variations of vowels following these consonants. (Abramson and Erickson 1992; Abramson 1997; Haudricourt 1956; Matisoff 1973; Ohala 1973; and

* The author would like to thank Professor John J. Ohala for his advice and comments and Dr. Natasha Lynn Warner for her help with the statistical analysis.

TAI STUDIES PROCEEDINGS 361-369

Whalen et al. 1993) Vowels after original voiceless stops have a higher tone; whereas vowels after voiced stops have a lower tone. The development of these Fo distinctions led to the split of tones in Tai languages from three tones in Proto Tai to six tones in Lungchow and Po-ai and five tones in Siamese. The tones, and examples from each language, can be seen below (adapted from Li 1977 and Abramson and Erickson 1992):

Lungchow: six tones: 3-3, 3-1, 5-5, 1-1, 2-4, and 2-1.

| Proto-Tones[1] | [voice] | Tones | Examples |
|---|---|---|---|
| A | voiceless | 3-3 | *pai* 'to go' |
| | Voiced | 3-1 | *taa* 'to smear' |
| B | voiceless | 5-5 | *khau* 'knee' |
| | Voiced | 1-1 | *taa* 'wharf' |
| C | voiceless | 2-4 | *khaa* 'to kill' |
| | Voiced | 2-1 | *nam?* 'water' |

Po-ai: six tones: 2-4, 3-1, 5-5, 2-2, 3-3, and 4-4.

| Proto-Tones | [voice] | Tones | Examples |
|---|---|---|---|
| A | voiceless | 2-4 | *pai* 'to go' |
| | | or 3-1 | *?au* 'to take' |
| | voiced | 5-5 | *taa* 'to smear' |
| B | voiceless | 2-2 | *hoo* 'knee' |
| | Voiced | 3-1 | *taa* 'wharf' |
| C | voiceless | 4-4 | *kaa* 'to kill' |
| | Voiced | 3-3 | *lam* 'water' |

Siamese: five tones: 2-4, 3-3, 2-2, 4-1, and 4-5-3 (There is a merge between B2 and C1).

[1]Proto tones A, B, and C occur in live syllables (syllables which do not have stops as finals).

| Proto-Tones | [voice] | Tones | Examples |
|---|---|---|---|
| A | voiceless | 2-4 | *khaai* 'to sell' |
| | | or 3-3 | *pai* 'to go' |
| | voiced | 3-3 | *thaa* 'to smear' |
| B | voiceless | 2-2 | *khau* 'knee' |
| | Voiced | 4-1 | *thaa* 'wharf' |
| C | voiceless | 4-1 | *khaa* 'to kill' |
| | Voiced | 4-5-3 | *thɔɔŋ* 'stomach' |

Today, Tai languages such as Lungchow and Siamese have three sets of stops; voiceless aspirated stops (from Proto-Tai voiceless aspirated stops in Lungchow and Siamese, and Proto-Tai voiced stops in Siamese), voiceless unaspirated stops (from Proto-Tai voiceless unaspirated stops in Lungchow and Siamese, and Proto-Tai voiced stops in Lungchow), and voiced stops (from Proto glottalized voiced stops in Lungchow and Siamese). Some examples follow (from Li 1977):

| Proto Tai | Siamese | Lungchow | Tone |
|---|---|---|---|
| 1. Proto-Tai voiceless unaspirated stop (labial) | | | |
| **p* | p | p | |
| ex. 'year' | pii | pii | A1 |
| 2. Proto-Tai voiceless aspirated stop (labial) | | | |
| **ph* | ph | ph, p | |
| ex. 'male' | phuu | phuu | C2 |
| 3. Proto-Tai voiced stop (labial) | | | |
| **b* | ph | p | |
| ex. 'fat' | phii | pii | A2 |
| 4. Proto-Tai pre-glottalized stop (labial) | | | |
| **ʔb* | b | b | |
| ex. 'shoulder' | baa | baa | B1 |

It is of interest to see how important these Fo cues (or any other cues associated with the stops) are to Tai listeners and it was therefore decided to see whether an

experiment in the lab could duplicate one stage of the tonogenesis process in Tai. For this reason, a perceptual experiment was conducted.

## Experiment

The Siamese (Standard Thai) language was chosen for this experiment. The goal of this experiment was to see whether Siamese (Thai) listeners could differentiate voiced-voiceless stop minimal pairs without the main cues for the voicing itself.

Procedure. Six minimal pairs of Siamese (Standard-Thai) words were chosen for this experiment:

| Voiceless | meaning | voiced | meaning | tone |
|---|---|---|---|---|
| 1) paan | 'as if' | baan | 'to bloom' | [mid] |
| 2) pàa | 'forests' | bàa | 'shoulders' | [low] |
| 3) pâa | 'aunt' | bâa | 'crazy' | [rising-falling] |
| 4) tææŋ | 'melon' | dææŋ | 'red' | [mid] |
| 5) tàp | 'liver | dàp | 'to extinguish' | [low] |
| 6) tâan | 'to resist' | dâan | 'sides' | [rising-falling] |

Each pair contained a word with a voiceless unaspirated stop and the other, a word with a voiced stop.[2] The first three pairs had initial labial stops and the other three pairs had initial dental stops. Pairs 1 and 4 had a mid tone, pairs 2 and 5 had a low tone, and pairs 3 and 6 had a rising-falling tone. All the words were read in the context, "*nii* + a targeted word" meaning "This is____," recorded on an analog tape recorder, and subsequently digitised at a sample rate of 10 kHz.

In this experiment, we prepared 13 sets of tokens. The first set contained the unprocessed tokens which were the same as what we recorded (*nîi* + a targeted word). For the following six sets which were in the context *nîi* + a targeted word, we digitally gated the initial portions of each targeted word with six 25-ms steps. For the first step,

---

[2]I assume that the "voiced stops" in contemporary Central Thai (historically from the preglottalized series **ʔb* and **ʔd*) can serve to represent the "voiced stop," and the voiceless stops in contemporary Central Thai (historically from the voiceless stop series **p, *t, *k,* etc.) can represent the "voiceless unaspirated stops" for the purposes of this perceptual experiment, cf. Li 1977.

each token contained the context *nîi* and the targeted word from 30 ms *before* the stop onset to the rest of the word. For the second step, each token contained the context *nîi* and the target word from the stop onset to the rest of the word. For the third step, each token contained the context *nîi* and the targeted word from 30 ms *after* the stop onset to the rest of the word. For the fourth step, each token contained the context *nîi* and the targeted word from 60ms *after* the stop onset to the rest of the word. For the fifth step, each token contained the context *nîi* and the targeted word from 90 ms *after* the stop onset to the rest of the word. For the sixth step, each token contained the context *nîi* and the targeted word from 120 ms *after* the stop onset to the rest of the word. For the other six sets, we repeated the same six steps as above but we replaced the context *nîi* with a white-noise context. All the tokens were randomised.

There were a total of 156 randomised tokens (12 words (6 pairs) *6-step gating *2 contexts + a set of 12 unprocessed tokens [*nîi* + word]). Eleven adult native Thai subjects (six males and five females with ages ranging from 18-35) listened to these tokens and judged whether each token had a voiceless or voiced initial stop (ป /p/ or บ /b/; ต /t/ or ด /d/).

**Results**

Figures 1 and 2 show the results of this experiment. Both figures show the percentage of voiceless (in solid line) and voiced (in dashed line) responses over different gating points. The difference between the two figures was the contexts, namely, an onset word /nîi/ in figure 1 and a white-noise context in figure 2. From the two figures, it can be seen that listeners heard the voiceless tokens as voiceless more often than they heard the voiced tokens as voiced. When the gates were greater than 30 ms after the stop onset, although listeners began to increase their judgements for tokens they heard (from voiced and voiceless tokens) as voiceless, they were still able to distinguish between the voiceless and the voiced tokens at a higher rate than would be possible if only the voicing in stop closure were the cue. In other words, if the voicing in the closure were the only cue, all the judgements for voiced stops should have been 'voiceless' at gates + 30 ms or more.

**Analysis**

An analysis of variance (ANOVA) was used to analyse the interactions between voicing, places of articulations, and tonal effects of this experiment.

Firstly, voicing[3] had a significant effect on the responses at every point of time (30 ms before the stop onset, at the burst, 30 ms, 60 ms, 90 ms and 120 ms after the stop onset) (see #1). Listeners heard the voiceless tokens as voiceless more often than they heard the voiced tokens as voiced - as can be seen from the two figures, the correct responses of the voiceless tokens are higher than the correct responses of the voiced tokens. It is, however, interesting to note that listeners were able to differentiate between voiceless and voiced tokens although the cues for voicing past 30 ms after the stop onset were gated out (see the correct responses at 60 ms, 90 ms and 120 ms after the burst in figures 1 and 2).

Secondly, places[4] of articulation had a significant effect on responses at the stop onset, 30 ms, 60 ms, 90 ms, and 120 ms after the stop onset (see #2). We found that listeners heard voicing differently more for alveolars than for labials at these points in time.

Thirdly, tones[5] had a significant effect on responses at the stop onset and 120 ms after the stop onset (see #3). We found that listeners heard voicing differently for all tones (mid, low, and rising-falling) at the stop onset and for only the contour tone at 120 ms after the stop onset.

---

[3]#1. Voicing

| | | | |
|---|---|---|---|
| overall: | $F(1,10)= 495.87$, $p<0.0001$ | 30ms after: | $F(1,10)= 81.67$, $p<0.0001$ |
| 30ms before: | $F(1,10)= 1024.00$, $p<0.0001$ | 60ms after: | $F(1,10)= 36.75$, $p<0.0005$ |
| at the burst: | $F(1,10)= 337.84$, $p<0.0001$ | 90ms after | $F(1,10)= 45.00$, $p<0.0001$ |
| | | 120ms after: | $F(1,10)= 12.00$, $p<0.01$ |

[4]#2. Places

| | | | |
|---|---|---|---|
| overall: | $F(1,10)= 24.95$, $p<0.001$ | 60ms after: | $F(1,10)= 12.53$, $p<0.01$ |
| at the burst: | $F(1,10)= 10.20$, $p<0.01$ | 90ms after | $F(1,10)= 15.94$, $p<0.01$ |
| 30ms after: | $F(1,10)= 33.14$, $p<0.0005$ | 120ms after: | $F(1,10)<1$ |

[5]#3. Tones

| | | | |
|---|---|---|---|
| overall: | $F(2,20)= 7.32$, $p<0.005$ | 120ms after: | $F(1,10)= 8.33$, $p<0.02$ |
| at the burst: | $F(2,20)= 17.89$, $p<0.0001$ | | |

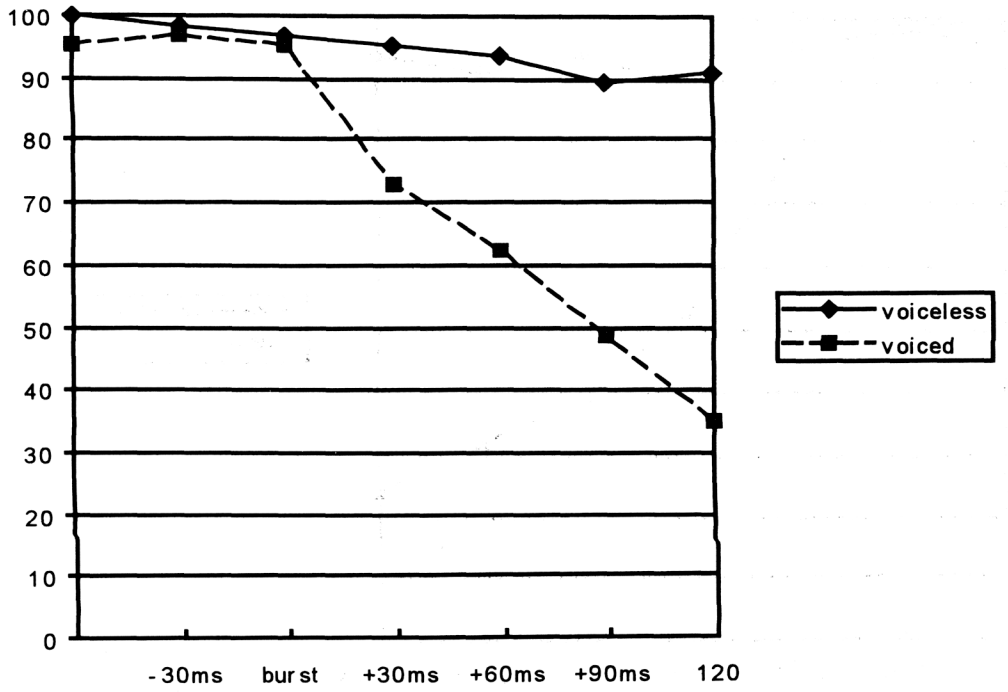


*Figure 1.* Correct responses: *nîi* context

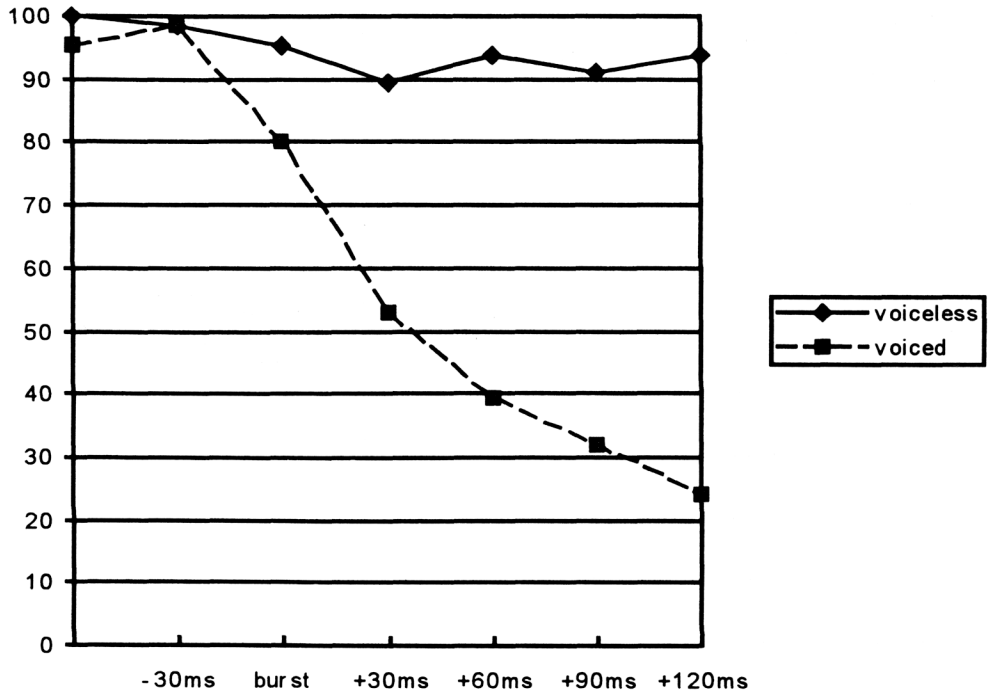


*Figure 2.* Correct responses: White noise context

**Discussion**

The data show that listeners heard the voiceless tokens as voiceless more often than they heard the voiced tokens as voiced. However, when the gates were greater than 30 ms after the stop onset, listeners were still able to distinguish between the voiceless and voiced tokens at a higher rate than would be possible if only the physical cue of voicing in stop closure were the cue. If Fo is one of the major cues besides the voicing in stop closure, it will confirm the theory of Whalen et al. (1993) as quoted "... Still, it is clear that the Fo value of the syllables as a whole, not just the Fo perturbation, can affect the voicing judgement..." This experiment also confirmed the findings of Abramson and Erickson (1992) that fundamental frequency can affect the perception of voicing.

This experiment, then, illuminates the sound change in Tai as represented by the Siamese language and, as mentioned above: the development of tones due to the merger of voiced and voiceless stops. However, there will need to be more studies to examine

the importance of Fo. If Fo turns out to be a major cue, then this experiment partially duplicates tonogenesis in Tai.

## REFERENCES

Abramson, Arthur S. and Donna Erickson. 1992. "Tone splits and voicing shifts in Thai: phonetic plausibility." *Pan Asiatic Linguistics: Proceedings of the Fourth International Symposium on Languages and Linguistics* 1:1-15.

Abramson, Arthur S. 1997. "The Thai tonal space." In Arthur Abramson ed., *Southeast Asian Linguistic Studies: In Honor of Vichin Panupong*. Bangkok: Chulalongkorn University Press.

Haudricourt, Andre. 1956. "De la restitution des initiales dans les langues monosyllabiques: Le probleme du Thai commun." *BSLP* 52:307-322.

Li, F.K. 1977. *The Handbook of Comparative Tai*. Honolulu: University of Hawaii Press.

Matisoff, James A. 1973. "Tonogenesis in Southeast Asia." In Larry M. Hyman ed., *CTT*.

Ohala, John J. 1973. "The physiology of tone." In Larry M. Hyman, ed., *CTT*.

Whalen, D.H. et al. 1993. "Fo gives voicing information even with unambiguous VOT." *JASA* 93(4):2152-2159.