Temporal Relations between Thai Initial Stops and Vowels: Acoustic and Perceptual Studies

Chutamanee Onsuwan
University of Michigan

1. Introduction

This study continues a line of inquiry begun by Onsuwan and Beddor (1998), examining possible acoustic and perceptual interactions between two temporal properties in Thai, vowel duration and stop consonant voice onset time (VOT). Although previous studies (e.g., Abramson and Ren 1990; Lisker and Abramson 1970) have investigated the phonetic characteristics of these two properties, the focus of the present study is the temporal relations that hold between the two.

Thai contrasts long and short vowels. All 9 vowels /i e ɛ a ɯ ɤ u o ɔ/ occur phonemically short and long. Thai initial stops show a 3-way contrast in voicing and aspiration (voiced, voiceless unaspirated and voiceless aspirated) in bilabial /b p pʰ/ and alveolar stops /d t tʰ/ and a 2-way contrast (voiceless unaspirated and voiceless aspirated) in velar stops /k kʰ/.

VOT and vowel length are both temporal distinctions. VOT is the time interval between obstruent release and the onset of voicing for a following vowel. Measurement of VOT and vowel duration is shown in Figure 1. In the waveform display (bottom panel), the release burst marks the beginning of VOT (marked by the first cursor). The onset of periodic pulsing marks the end of VOT as well as the onset of the vowel (marked by the second cursor). The vowel offset (marked by the third cursor) is identified by a change in the patterns of the periodic wave and by the end of energy in the lower formants shown in the wideband spectrographic display (top panel).

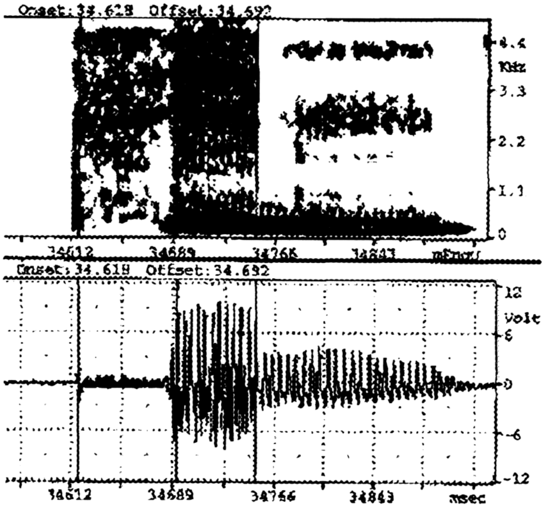

Figure 1. Wideband spectrogram (top panel) and waveform display (bottom panel) with the first cursor placed at the beginning of VOT (consonantal release burst); the second cursor at the end of VOT/vowel onset and the third cursor at the vowel offset. The word is [t^{h}ìn] 'region' produced by a female speaker. VOT measurement is 74.7 ms. Vowel duration is 60 ms.

However, during aspiration, the vocal tract is already in the position for the vowel. Aspiration can be viewed therefore as a voiceless articulation of vowel onset. For this reason, it is possible that, although consonant laryngeal features and vowel length are phonologically independent in Thai, phonetically vowel length might influence VOT and/or vice versa.

The possible effect of phonological consonant laryngeal timing on phonetic vowel duration and of phonological vowel length on phonetic VOT is explored here in terms of acoustic measures of these temporal relations and perceptual study of the effect of these relations on listeners' judgments. The study involves 2 parts. The first part investigates possible influences of the phonological contrast between aspirated and unaspirated stops on the phonetic duration of following vowels. The results will show that the aspirated-unaspirated contrast does affect following vowel duration both acoustically and perceptually. The second part of this study investigates whether phonological contrasts in vowel length affect phonetic voice onset time of preceding stops. The results will show that although acoustically phonological vowel length has no effect on stop VOT, Thai listeners' perception of VOT is proportional to following vowel length.

2. *Part 1:* Influences of the Phonological Contrast between Aspirated and Unaspirated Stops on the Phonetic Duration of Following Vowels

2.1 Acoustic Analysis

2.1.1 *Method*

Three native Thai speakers participated in the acoustic study. The speakers were recorded reading multiple repetitions of a list of 40 minimal or near-minimal pairs. The members of each pair began with a voiceless unaspirated or aspirated stop and differed from each other in terms of vowel length (e.g., [pān]-[pāan] 'to share-birthmark'; [pʰān]-[pʰāan] '1000-tray with pedestal'; [pàt]-[pàat] 'to brush way-to smooth and level off'; [pʰàt]-[pʰàat] 'to stir fry-briefly'; [tʰōn]-[tʰōon] 'to bear-singleton'; [kàk]-[kàak] 'to stop-garbage'). Test pairs were balanced as evenly as possible, given the constraints imposed by using real-word pairs, across 3 places of articulation (bilabial, alveolar, and velar), 5 vowels (/i e ɛ a o/) and 3 level tones (high, mid, and low).

Five repetitions of the 80 words were randomized for a total of 400 items for each speaker. Each item was read in a sentence context [ʔàan wâa...ʔìik] 'read as__again'. The recorded materials were then digitized and analyzed acoustically in terms of 2 temporal measures: VOT and vowel duration. VOT duration is measured from the release burst until the onset of periodic pulsing which also marks the onset of the vowel. Vowel duration is measured from the vowel onset until the vowel offset generally identified by the end of energy in the lower formants.

2.1.2 *Results*

The vowel duration measures are shown in Figure 2. It can be seen from this figure that the duration of long vowels was approximately twice as long as that of short vowels, and that all vowel qualities showed roughly the same relation.

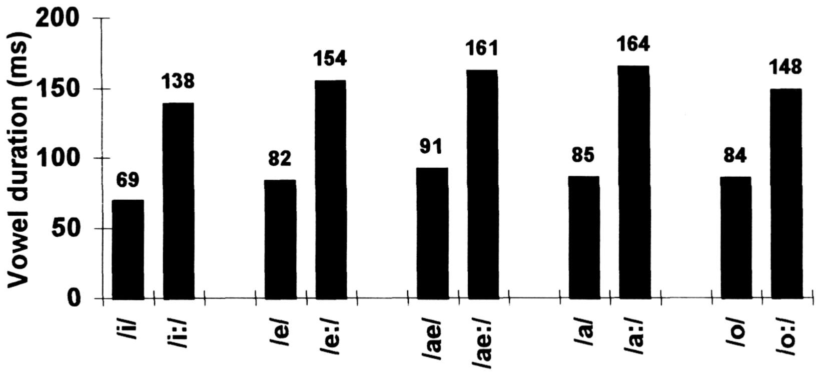

Figure 2. Duration (in ms) of long and short vowels pooled across 3 speakers and consonantal contexts[1].

The main question here however is whether the phonetic vowel duration in the context of aspirated stops differs from that in the context of unaspirated stops. Figure 3 shows that the duration of the voiced portion of both short and long vowels (the dark portion of each bar in this figure) was shorter in the context of aspirated stops than in the context of unaspirated stops. This pattern holds for all places of articulation. Short vowels were on average 21% shorter in the context of aspirated stops and long vowels were on average 9% shorter.

Alternatively, if aspiration is viewed as voiceless portion of following vowel, vowel duration would correspond to VOT plus the following voiced vocalic portion (in this figure, each entire bar would correspond to vowel duration). Therefore, the alternative account for the pattern would be that phonetic vowel duration is longer preceded by aspirated consonants than by unaspirated consonants. However, in the case of vowels preceded by aspirated consonants, proportionally less of the vowels is voiced.

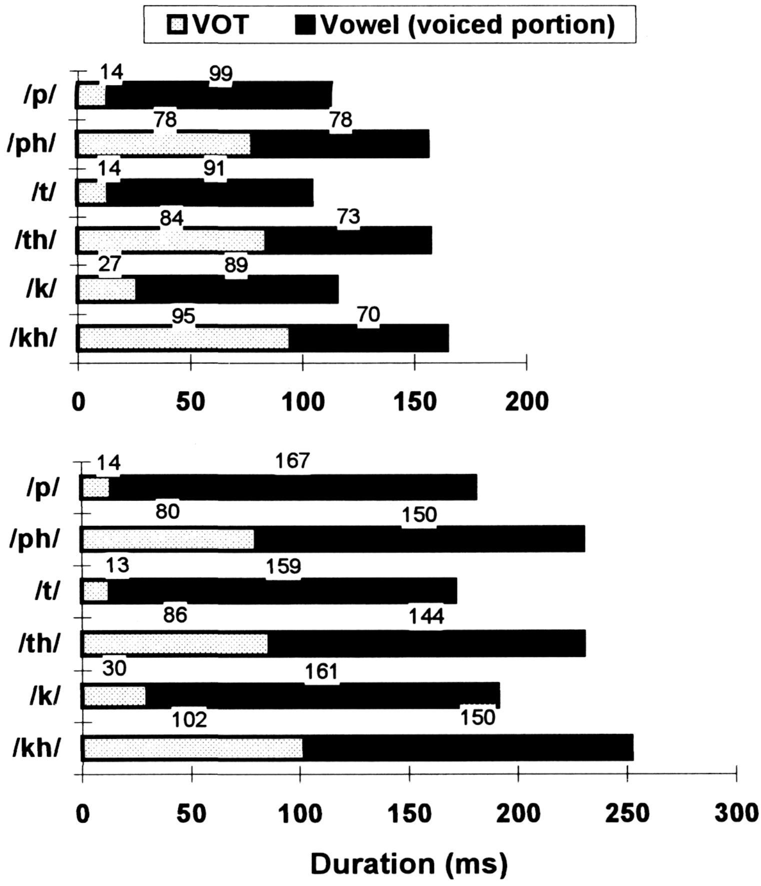

Figure 3. Average duration in millisecond of VOT plus short vowels (top panel), and VOT plus long vowels (bottom panel) according to places of articulation and aspiration. Measures pooled across 3 speakers.

In summary, results from the acoustic study showed that the aspirated-unaspirated contrast had an effect on following vowel duration, with the voiced portion of vowels being shorter after aspirated stops. However, if aspiration is viewed as a voiceless portion of the vowel, the total vowel portion is longer after aspirated than after unaspirated stops.

One way of viewing this relation is that aspiration adds to total syllable length and that the shortened voiced portion of the vowel after aspirated stops is compensatory shortening. This raises the question of how sensitive listeners are to this timing relation.

2.2 Perceptual Study

The first perception experiment investigated a possible effect of the aspiration contrast on the perception of vowel length.

The acoustic study showed that the aspiration contrast had an effect on the phonetic duration of the following vowel, such that the voiced portion of vowels is shorter after aspirated stops. Given the acoustic effect of aspiration on following vowel duration, it is expected that perceptually whether the initial stop is aspirated or unaspirated will affect the perception of length of the following vowel. More specifically, after an aspirated stop, it is predicted that listeners would need less (phonetic) vowel duration to perceive a vowel as long.

2.2.1 *Method*

A 17-step vowel duration continuum ranging from 86-180 milliseconds (with 4-6 milliseconds increments) was created by excising and editing the vowel portion of a natural token [pàat] produced by a female speaker. The vowel continuum was then spliced onto the VOT portions of 2 natural tokens: [pàat] and [pʰàat]. Altogether, there were 2 vowel series, one with unaspirated portion: [pàt]-[pàat] and the other with aspirated portion: [pʰàt]-[pʰàat].

Twelve repetitions of the 17-step vowel stimuli (204 trials per vowel series) were randomized and presented to 15 Thai listeners for identification. They were asked to identify the vowel in each stimulus as short /a/ or long /aa/.

2.2.2 *Results*

The question here is whether identification of vowels as short or long shifted as a function of preceding aspiration, as expected on the basis of the acoustic data. Figure 4 gives the percent short vowel responses to the vowel continuum in unaspirated versus aspirated contexts. As expected, the identification results showed that the vowel continuum preceded by an aspirated stop elicited fewer short vowel responses than that preceded by an unaspirated stop. A possible explanation comes from the acoustic findings, where the voiced portion of vowels (either long or short) is shorter if preceded by aspirated stops. It is expected that Thai listeners are aware of this relation. To identify a vowel of a given duration which preceded by stop consonants, listeners make use of the VOT-to-vowel relation and require less vowel duration with initial aspirated stops in order to label the vowel as long.

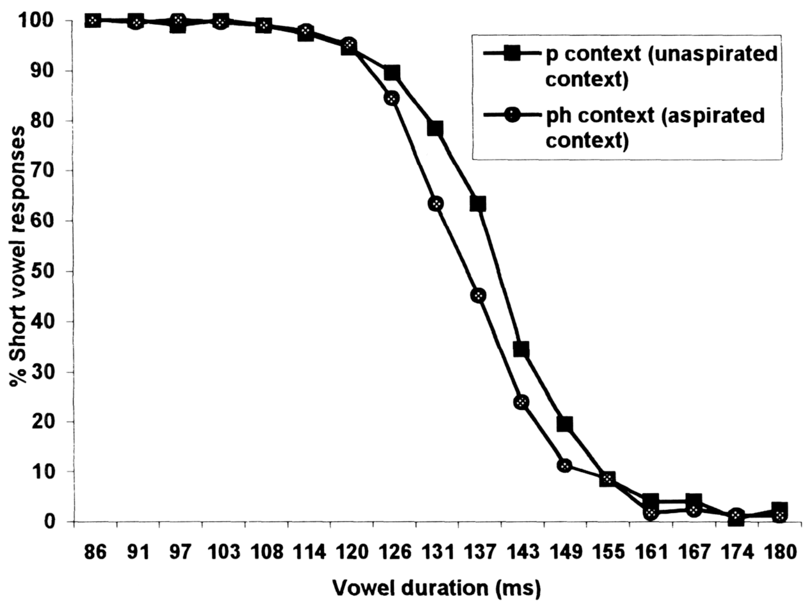

Figure 4. Pooled identification responses of 15 Thai listeners to 17-step vowel continuum preceded by unaspirated (squares) and aspirated (circles) stop.

3. *Part 2:* Influences of the Phonological Contrast in Vowel Length on the Phonetic VOT of Preceding Stops

3.1 Acoustic Analysis

3.1.1 *Method* (Same method as for Part 1)

Three native Thai speakers participated in the acoustic study. The speakers were recorded reading multiple repetitions of a list of 40 minimal or near-minimal pairs. The members of each pair began with a voiceless unaspirated or aspirated stop and differed from each other in terms of vowel length (e.g., [pān]-[pāan] 'to share-birthmark'; [pʰān]-[pʰāan] '1000-tray with pedestal'). Test pairs were balanced as evenly as possible, given the constraints imposed by using real-word pairs, across 3 places of articulation (bilabial, alveolar, and velar), 5 vowels (/i e ɛ a o/) and 3 level tones (high, mid, and low).

Five repetitions of the 80 words were randomized for a total of 400 items for each speaker. Each item was read in a sentence context [ʔàan wâa...ʔìik] 'read as__again'. The recorded materials were then digitized and analyzed acoustically in terms of 2 temporal measures: VOT and vowel duration. VOT duration is measured from the release burst until the onset of periodic pulsing which also marks the onset of the vowel. Vowel duration is measured from the vowel onset until the vowel offset generally identified by the end of energy in the lower formants.

3.1.2 *Results*

The main question here is whether phonetic VOT duration in the context of phonologically short vowels differs from that in the context of phonologically long vowels. The results of the VOT measures are shown in Figure 5 and 6. For the unaspirated stops (Figure 5), across the 3 places of articulation, mean VOT of unaspirated stops was 19 milliseconds when followed by short vowels and 20 milliseconds when followed by long vowels. No place of articulation showed a significant effect of distinctive vowel length on VOT. For the aspirated stops (Figure 6), mean VOT was 87 milliseconds when followed by short vowels and 91 milliseconds when followed by long vowels. For these stops, analyses of each place of articulation showed that VOT duration did not vary significantly as a function of phonological vowel length.

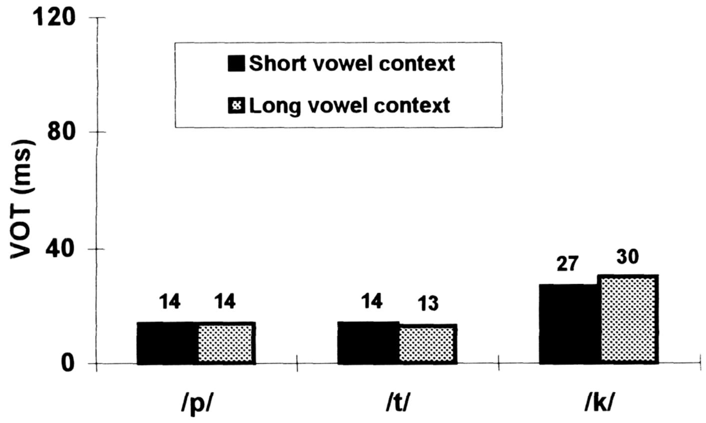

Figure 5. VOT (in ms) of unaspirated stops in short and long vowel contexts. Measures pooled across 3 speakers.

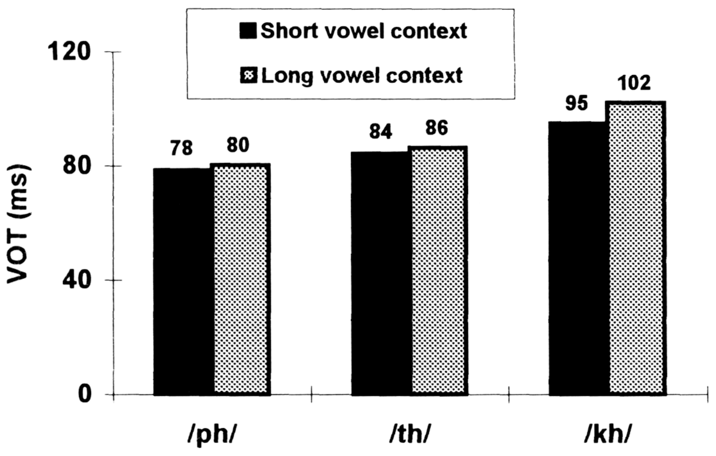

Figure 6. VOT (in ms) of aspirated stops in short and long vowel contexts. Measures pooled across 3 speakers.

In summary, acoustically, the distinction between phonological short and long vowels did not have any effect on phonetic duration of VOT.

3.2 Perceptual Study

Although the findings from the acoustic study showed that acoustically phonological vowel length did not affect phonetic duration of VOT, it is possible that Thai listeners' perception of VOT might be influenced by the length of the following vowel. As a native speaker of Thai, I hear the VOT of stop consonants followed by short vowels to be longer than that followed by long vowels. The perception on my part may be related to the fact that long vowels are generally twice as long as short vowels in Thai. Therefore, a given amount of aspiration might sound proportionally shorter before a long vowel than it would if it were near a short vowel.

The second perception experiment was conducted to investigate possible effects of phonological vowel length on the perception of VOT. The main question is whether the way Thai listeners perceive VOT in the context of short vowels differs significantly from that in the context of long vowels. If perception of VOT is proportional to vowel length, Thai listeners should require more aspiration to hear a stop as aspirated when followed by a long vowel than when followed by a short vowel.

3.2.1 *Method*

A 13-step VOT continuum ranging from 0-70 milliseconds was created by editing the aspiration of naturally produced [phān] (produced by the same female speaker as in the first perception experiment). Aspiration increments were 5 milliseconds in the 0-50 milliseconds range and 10 milliseconds in the 60-70 milliseconds range.

The VOT continuum was then spliced onto 2 different vowel portions, a long vowel and a short vowel. I began with the long vowel from a natural token [pāan]. The new, edited short vowel portion [pān] was created from the natural token. The long vowel [pāan] was shortened by deleting successive pitch pulses (vowel duration of the original long was 180 milliseconds, and the shortened long 86 milliseconds). By using edited shortened version of the original vowel portion, the goal is to hold all spectral properties of the vowel constant and to vary only temporal properties.

The 13-step VOT continuum was spliced onto the 2 vowel portions creating 2 VOT series: one with long vowel from the original long vowel and the other with short vowel from the shortened long vowel. Twelve repetitions of the 13-step VOT stimuli (156 trials per VOT series) were randomized and presented to 14 Thai listeners for identification as beginning with either aspirated [pʰ] or unaspirated [p].

3.2.2 *Results*

The main question here is whether identification responses to stimuli as aspirated or unaspirated shifted as a function of the phonological length of the following vowel or whether the 50% boundary crossover points for both series differed.

Figure 7 gives the percent unaspirated stop responses in short versus long vowel contexts.

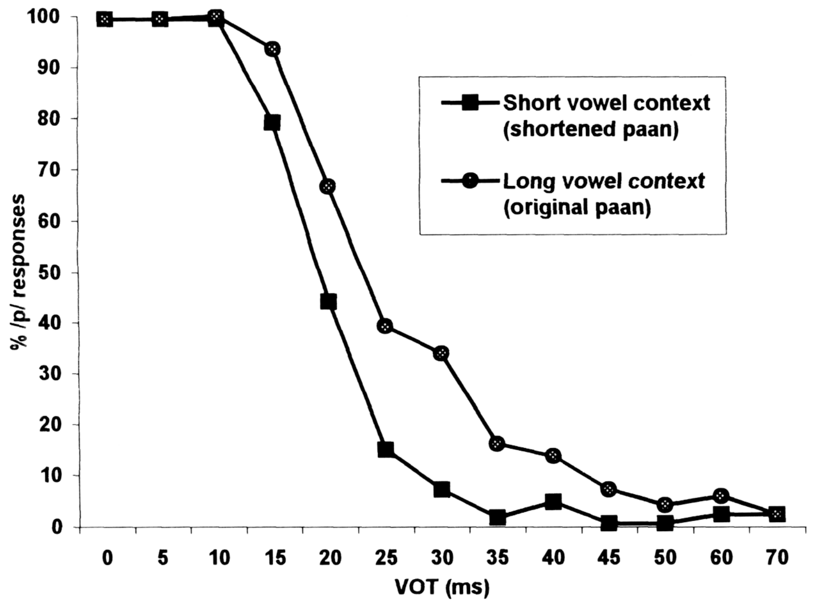

Figure 7. Pooled identification responses of 14 Thai listeners to 13-step VOT continuum in short (squares) and long (circles) vowel contexts. Vowel portions excised from original [pāan].

As predicted, the identification results showed that the VOT continuum followed by the long vowel (the original long vowel) elicited more unaspirated responses (or fewer aspirated responses) than the continuum followed by the short vowel (the edited short vowel) did.

As expected if VOT perception is proportional to following vowel length, the results suggest that, in the middle of the VOT continuum where there is some ambiguity, listeners are more likely to label a stop as unaspirated if the following vowel is long. Phonological long vowels in Thai are generally twice as long as phonological short vowels, which is also true of the vowel tokens used in this perceptual study. As mentioned previously, next to a long vowel, the same amount of stop aspiration should sound shorter than it would if it were near a short vowel. Apparently, these proportional differences influenced listeners' perception of VOT duration and caused listeners to perceive stops more frequently as unaspirated before the long vowel.

4. Conclusion

In *Part 1*, where there was an acoustic effect of the phonological contrast between aspirated and unaspirated stops on following phonetic vowel duration, with the voiced portion of vowels being shorter after aspirated stops, results from the perception experiment showed that listeners made use of the acoustic VOT-to-vowel relation in identifying vowels as long or short. However, in Part 2, where there was no acoustic influence of phonological vowel length on stop VOT, the perception experiment showed that there was also a perceptual effect of vowel length contrast on VOT identification. My speculation was that in the second case, the perceptual influence despite the lack of an acoustic difference was due to the region of ambiguity. In this region, specifically, the proportional differences between aspiration and vowel length influenced listeners' perception of VOT.

Taken together, results from *Part 1* and *Part 2* suggest that the two temporal properties, VOT of initial stops and vowel length interact. Moreover, the perceptual experiments suggest that the temporal relations between VOT and vowel length serve as possible phonetic cues for Thai listeners in the perception of vowel length and aspiration contrast.

Note

1. In Figure 2, /ae/ is used to refer to /ɛ/, and /ae:/ to refer to /ɛ:/.

References

Abramson, Arthur S. and Ren, Nianqi. 1990. Distinctive vowel length: duration vs. spectrum in Thai. Journal of Phonetics 18. 79-92.

Lisker, Leigh and Abramson, Arthur S. 1970. The voicing dimension: some experiments in comparative phonetics. Proceedings of the 6th International Congress of Phonetic Sciences, Prague, 563-567.

Onsuwan, Chutamanee and Beddor, Patrice S. 1998. Acoustic and perceptual effects of vowel length on voice onset time in Thai stops. Proceedings of the 16th International Congress on Acoustics and the 135th Meeting Acoustical Society of America, Seattle, 2943-2944.