Durational and Spectral Differences in Thai Diphthongs and Final Glides

Phongphat Man-khongdi[1], Chutamanee Onsuwan[1,2], Charturong Tantibundhit[2,3]

[1]Department of English and Linguistics, Thammasat University, Thailand
[2] Center of Excellence in Intelligent Informatics, Speech and Language Technology and Service Innovation (CILS), Thammasat University, Thailand
[3]Department of Electrical and Computer Engineering, Thammasat University, Thailand

phtmankhongdi@gmail.com, consuwan@tu.ac.th, tchartur@engr.tu.ac.th

Abstract

Acoustic analysis was conducted to compare Thai monophthongs /i, ii, ɯ, ɯɯ, u, uu, a, aa/, diphthongs /ia, ɯa, ua/, and vowel-to-glides (vowel + /j/ or /w/) in terms of duration, formant frequency, and spectral rate of change (TL_{roc}). Preliminary results from multiple repetitions of 30 target monosyllabic words show not only that Thai diphthongs and vowel-to-glides differ in their articulatory trajectories, but they appear to differ in their duration and TL_{roc} values. Average duration of diphthongs is shorter than that of vowel-to-glides. TL_{roc} of diphthongs is on average higher than that of long-vowel-to-glides, but lower than short-vowel-to-glides.

Index Terms: acoustics, diphthongs, final glides, Thai

1. Introduction

Despite the fact that vowels and glides sound quite similarly, it is widely accepted that phonetically, vowels and glides are distinct from one another. Glides do not maintain a steady state, but vowels (including diphthongs) do [1] and [2]. Some acoustic differences have been noted between diphthongs and glides. Acoustically, diphthongs and glides are investigated by analyzing a transition portion of two target sounds. The transition of diphthongs is usually slower and more gradual than that of vowel-to-glides [3]. This could be due to the fact that the production of glides moves in and out very quickly. Kenyon and Jones (1924) [4] stated that "rapid movement establishes /j, w/ as a set of 'gliding consonants' resulting from an immediate rapid movement of the lips and the tongue, or tongue alone" [5]. Gimson (1962) [6] also supported this idea and added that "a semivowel is a rapid vocalic glide on to a syllabic sound of greater steady duration" [4].

Of interest here are the diphthongs and final glides of Thai. Phonologically, Thai has three diphthongs, /ia/, /ɯa/ and /ua/. Unlike the 9 monophthongs which are phonemically contrastive in length, the diphthongs are not. However, their durations are predictable when followed by different finals [7], [8], [9], and [10]. Phonetically, /ia/, /ɯa/ and /ua/ could all be classified as opening diphthongs, that is, the onset is one of high vowels (/i/, /ɯ/, /u/), and the offset is low vowel (/a/). It should be noted that /i/, /ɯ/, /u/ and /a/ are among Thai monophthongs with long vowel counterparts. As for Thai glides, there are /j/ and /w/, each occurs in word initials and finals. Compared with initial glides, the final glides exhibit more limited co-occurrences with the vowels [11].

Only a handful of phonetics studies have been devoted to Thai diphthongs and glides. Among them, Roengpitya [9] investigated Thai diphthongs, /ia/, /ɯa/ and /ua/ in terms of duration and found that duration of the onset was longer than the offset in all diphthongs; duration of the offset was a crucial acoustic cue leading to the short-long difference. Later, Roengpitya [10] also examined duration as well as F1 and F2 at the 25%, 50% and 75% time points of Thai diphthongs. She concluded that durational differences among the diphthongs were related to syllable types. Moreover, in syllables where diphthongs were followed by /w/, F2 values were likely to increase at the 75% point, and where diphthongs were followed by /j/, F2 values are likely to decrease at the 75% point. In those two studies, however, differences between the Thai diphthongs and glides were not explicitly compared.

Tingsabadh and Abramson noted that when following monophthongs or diphthongs, /j/ and /w/ sound similar to /i/ and /u/, respectively [8]. Especially, when each of the two glides follows the low front vowel /a/ or /aa/; such sequences are in opposite directions to Thai (opening) diphthongs. On this ground, it is interesting to investigate whether and to what extent such vowel-to-glides differ acoustically from the diphthongs, and which acoustic characteristics could possibly set them apart. Another interesting point to explore is concerning Thai monophthongs and diphthongs. As previously mentioned, the onset (/i/, /ɯ/, /u/) and offset (/a/) of the three diphthongs are inherently monophthongs, it is worth investigating how much /i/, /ɯ/, /u/ and /a/ as part of diphthong trajectories differ from the monophthongs.

Therefore, the goal of this study is twofold. Firstly, we examine an extent to which the onset (/i/, /ɯ/, /u/) and offset (/a/) as part of diphthongs differ from when they occur as monophthongs. Secondly, we explore acoustic correlates which could be accounted for differences among Thai diphthongs and vowel-to-glides. Three acoustic measures: duration, F1 and F2, and spectral rate of change (TL_{roc}) are taken.

In the following sections, experimental method, main findings and discussions are given.

2. Method

2.1. Participants

Three male speakers of Thai (age range, 24-26 years) participated in this study. All of them were born and raised in the central region of Thailand and use Bangkok Thai in their daily life. They reported no known speech or hearing disorders and presented no speech production problems.

2.2. Word list

There were 30 target words in total. All are meaningful monosyllabic Thai words, begin with voiceless stop and have level tone (mid, low, or high). The words could be classified into 9 groups according to syllable types as shown in Table 1.

Each word was embedded in a carrier sentence [kam laŋ phû:t *target word* ta:m pà ka ti] (I am speaking *target word* naturally) and read from a printout three times by each speaker. On the printout, the words were presented randomly in blocks. The last two repetitions of each were selected and analyzed.

Table 1: *30 target words (9 syllabic types) used in the study.*

Syllable types	Target words
Vʔ (short vowel with final stop)	tìʔ, pàʔ, ʔɯ̀ʔ, pùʔ, thɯ̀k
VV (long vowel with no final or final stop)	pii, paa, khɯɯ, tuu, pìik, pàak, pùut
Vw (short vowel with /w/)	tiw, taw
Vj (short vowel with /j/)	puj, kàj
VVw (long vowel with /w/)	taaw
VVj (long vowel with /j/)	taaj, kaaj
VD (diphthong with no final or with final stop)	pia, píaʔ, phɯ̀a, tua, ʔúaʔ, pìak, thɯ̀ak, pùat
VDw (diphthong with final /w/)	piaw
VDj (diphthong with final /j/)	pɯ̀aj, pùaj

2.3. Recording

All of the data were recorded by a recorder, MP3 Samsung model YP-Q2ABin a soundproof room at the Faculty of Liberal Arts, Thammasat University, Tha Phrachan Campus. Each recording session took about 10 minutes. The selected tokens were segmented and analyzed using PRAAT version 5.4.15 [12].

2.4. Measurements

Three measures were taken: duration of vowel (and final consonant), formant frequency (F1 and F2), and spectral rate of change (TL_{roc}) [13]. To obtain reliable comparisons among the vocalic portions (in ms) of monophthongs, diphthongs, and vowel-to-glides, final consonants (/w/, /j/, /t/, /k/ and /ʔ/) were included in the duration measure. It is worth noting that phonemically in Thai syllables with short vowel always ends with a final and that final stops are unreleased.

From the vocalic portion's duration, F1 and F2 values were manually extracted, at the 0%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% and 100% points (11 points in total). Then, spectral rate of change (TL_{roc}) was calculated based on [13]. Specifically, the length of vowel section (VSL) was calculated for each of five separate sections, i.e., 30%-40%, 40%-50%, 50%-60%, 60%-70%, and 70%-80% using the formula

$$VSL_n = \sqrt{(F1_n - F1_{n+1})^2 + (F2_n - F2_{n+1})^2}. \quad (1)$$

Then, overall formant TL defined as a summation of trajectories of five vowel sections was calculated by

$$TL = \sum_{i=1}^{5} VSL_n. \quad (2)$$

The TL_{roc} over the 50% portion of the vowel was calculated using

$$TL_{roc} = \frac{TL}{0.5V_{dur}}. \quad (3)$$

Finally, vowel section roc (VSL_{roc}) was calculated for each individual vowel section determined by the temporal location of the six measurement points, i.e., 30%-40%, 40%-50%, 50%-60%, 60%-70%, and 70%-80% using the formula

$$VSL_{roc_n} = \frac{VSL_n}{0.1V_{dur}}. \quad (4)$$

3. Results

The results are presented in terms of duration (section 3.1); values of F1 and F2 (section 3.2); and spectral rate of change (TL_{roc}) (section 3.3), as follows.

3.1. Duration

Figure 1 shows vocalic durations of the nine syllable types ranking from shortest to longest. Durational trend appears that VVw/j is the longest, followed by VDw/j, Vw/j, VV, VD, and Vʔ.

Durational ratio of long and short monophthongs is 2.2. Interestingly, average duration of diphthongs (VD = 227 ms) falls between, but is quite different from that of short and long monophthongs (Vʔ = 111 and VV = 248 ms).

For each diphthong, the average durations of /ia/, /ɯa/ and /ua/ are relatively comparable at 213.26, 228.49, and 239.32 ms respectively.

Short and long vowel-to-glides (Vw/j and VVw/j) are clearly longer than diphthongs (VD). Interestingly, when vowels are followed by glides, durational differences between short and long vowels become smaller and the whole portions (vowel + glide) are noticeably longer.

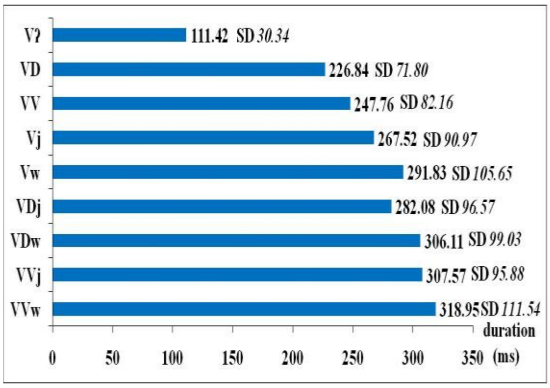

Figure 1: *Average duration of monophthongs and diphthongs, without any finals, or with final stops or final glides (Abbreviations as noted in Table 1).*

3.2. Formants

Figure 2 shows average F1 and F2 values for short and long monophthongs (at 50% point), diphthongs (solid lines showing downward trajectory at 0%, 50%, and 100% points), and vowel-to-glides /aj/, /aaj/, /aw/, and /aaw/ (dotted lines showing upward trajectory at 0%, 50%, and 100% points) and /iw/, /uj/ (dashed lines showing backing and fronting trajectories at 0%, 50%, and 100% points). Different types of vowel-to-glide are presented here to show an extent to which they are affected when preceded by low (dotted lines) versus high vowels (dashed lines).

It is clear that short versus long vowels of the same monophthong are relatively similar in terms of F1 and F2 (an exception might be for the /u/-/uu/ pair).For the three opening diphthongs, we could observe spectral changes in F1 and F2 values, specifically more so for F2 in the cases of /ua/ and /ia/. The diphthong onsets (/i/, /ɯ/, /u/ at 0%) and offsets (/a/ at 100%) as part of diphthongs appear to be more centralized than when they occur as monophthongs (exception might be for /ɯ/).

Similarly, onsets (/a/, /i/, /u/ at 0%) and offsets (/j/, /w/ at 100%) of the vowel-to-glides are relatively centralized.

3.3. TLroc

Spectral changes (movement) could be observed by F1-F2 plot (Figure 2), but TL_{roc} values certainly give us a clearer comparison when the length of vowel section is individually taken into account (See Section 2.4).

Figure 3 shows TL_{roc} values of the nine syllable types ranking from lowest to highest. The values seems to suggest the trend with VDw/j being the highest followed by Vw/j, VD, VVw/j, Vʔ, and VV.

As expected, short and long monophthongs have the lowest TL_{roc} values. TL_{roc} of diphthongs (VD = 5.5 Hz/ms) lies between, but is quite different from that of short vowels with final glide and long vowels with final glide (Vw/j = 6.2 and VVw/j = 3.6 Hz/ms).

Separately, TL_{roc} values of each diphthong vary from /ɯa/ (2.49 Hz/ms), /ua/ (6.36 Hz/ms) to /ia/ (6.66 Hz/ms).

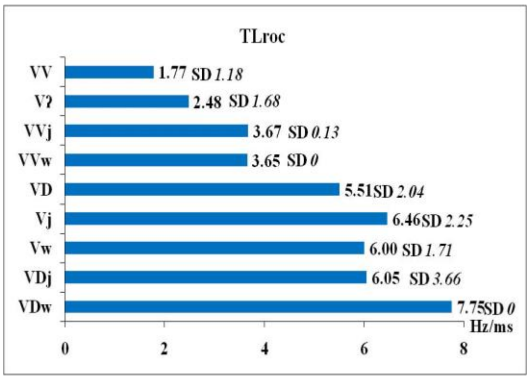

Figure 3: *TLroc values of monophthongs and diphthongs, without any finals, or with final stops or final glides (Abbreviations as noted in Table 1).*

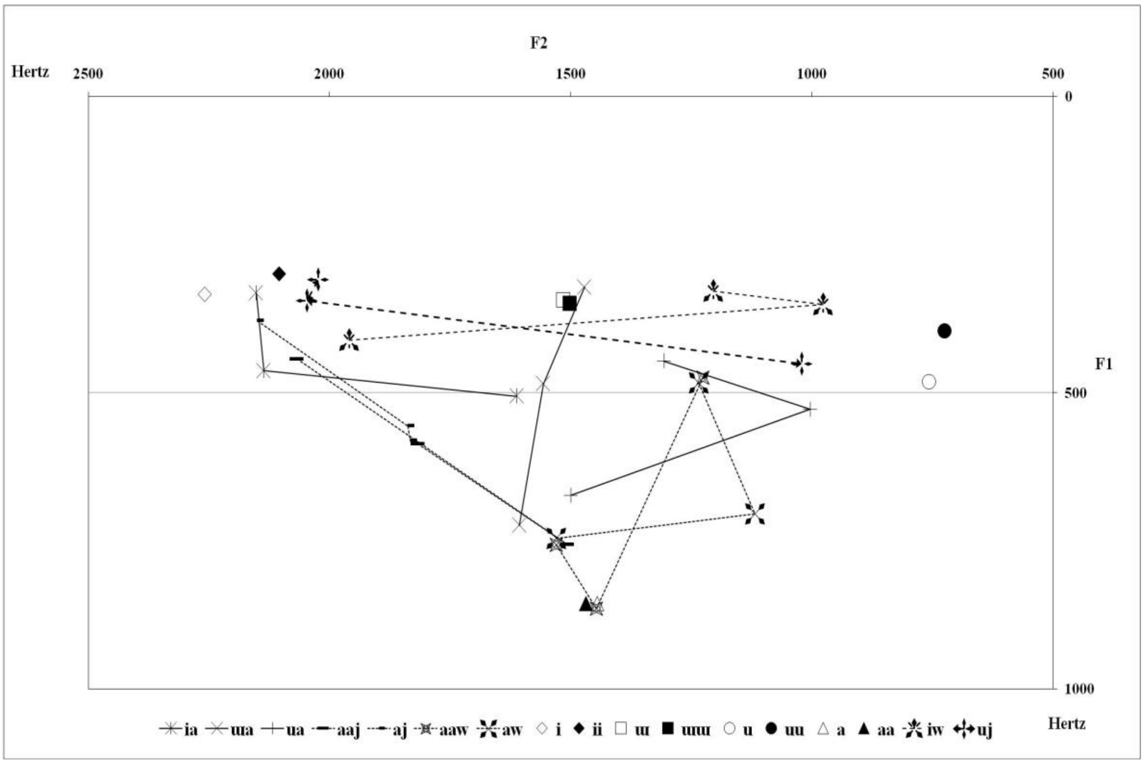

Figure 2: *Average F1 and F2 values of short and long monophthongs (at 50% point), diphthongs (solid lines showing downward trajectory at 0%, 50%, and 100% points), and vowel-to-glides /aj/, /aaj/, /aw/, /aaw/ (dotted lines showing upward trajectory at 0%, 50%, and 100% points) and /iw/, /uj/ (dashed lines showing backing and fronting trajectories at 0%, 50%, and 100% points).*

4. Discussions and future work

Durational values of short and long monophthongs as well as F1 and F2 range reported here are in agreement with previous studies of Thai [7]. The findings show that for the most part Thai diphthongs and vowel-to-glides are relatively different in terms of duration, F1 and F2 values, and spectral rate of change. Principally, in the case of Thai, differences in duration and spectral rate of change seem to be more relevant. Diphthongs are relatively shorter than either short or long monophthongs with glide. As noted by Roengpitya [9] and [10], Thai syllables with final glide appear to be longer than with any other types of final consonant (nasals and stops) and durational differences between short and long vowels become much less. This seems to be the case here, although direct comparison among different syllable types with various types of final consonant could provide a clearer picture.

In terms of F1 and F2 values for the Thai diphthongs, F2 dimension seems to capture spectral changes more clearly than F1, particularly for /ua/ and /ia/. Moreover, the diphthong onsets (/i/, /ɯ/, /u/) and offsets (/a/) as part of diphthongs appear to be more centralized than when they occur in monophthongs.

Lastly, TL_{roc} values, which have been used in a few studies of diphthongs and glides of other languages [13], [14] and [15], seem to be useful and reliable for capturing spectral dynamics of diphthongs and vowel-to-glides in this current study. TL_{roc} of Thai diphthongs is found on average to be higher than that of long-vowel-to-glides, but lower than short-vowel-to-glides. It remains to be seen with larger acoustic data if our finding agrees with the statement saying that the transition of diphthongs is usually slower and more gradual than that of the vowel-to-glide [3].

It is noteworthy that degrees of intensity were included in some studies of diphthongs and glides [16] and [17]. However, our preliminary acoustic analysis failed to show any noticeable differences in intensity level among Thai diphthongs and glides.

Finally, our future direction is to further explore acoustic properties of Thai diphthongs and glides from more speakers and to conduct detailed statistical analysis.

5. References

[1] Kent, R. D. and Read, C. *The acoustic analysis of speech* (2 ed.). Singular Publishing Group, 2002.

[2] Raphael, L. J., Borden, G. J. and Harris, K. S. *Speech science primer: physiology, acoustics, and perception of speech* (5 ed.). Lippincott Williams and Wilkins, 2006.

[3] Olive, J. P., Greenwood, A. and Coleman, J. *Acoustics of American English Speech: a dynamic approach*. New York: Springer-Verlag, 1993.

[4] Kenyon, J. S. In: Lance, D.M., Kingsbury, S.D. (Eds.), *American Pronunciation*. George Wahr Publishing Company, Ann Arbor, MI, 1924. (Citation from expanded 12th edition, 1997).

[5] Maddieson, I. "Glides and germination," *Lingua*, 118, 1926-1936, 2008.

[6] Gimson, A. C. *An Introduction to the Pronunciation of English*. London: Edward Arnold, 1962.

[7] Abramson, A. S. *"The vowels and tones of standard Thai: Acoustical measurements and experiments*," Indiana U. Research Center in Anthropology, Folklore, and Linguistics, Pub. 20. Bloomington, 1962.

[8] Tingsabadh, K. and Abramson, A. S. "Thai," *Journal of the International Phonetic Association*, 23:1, 1993.

[9] Roengpitya, R. "A Study of Vowels, Diphthongs, and Tones in Thai." Ph.D. dissertation. University of California at Berkeley, 2001.

[10] ___________. "Different Durations of Diphthongs in Thai: a New Finding," *The Proceedings of the Annual Meeting of Berkeley Linguistics Society*, 2002.

[11] Munthuli, A., Sirimujalin, P., Tantibundhit, C., Kosawat, K. and Onsuwan, C. "A Corpus-based Study of Phoneme distribution in Thai," *Proceedings of the 10th International Symposium Natural Language Processing (SNLP), Phuket, Thailand*. 114-121, 2013.

[12] Boersma, P. and Weenink, D. "Praat: doing phonetics by computer [Computer program]. Version 5.4.15," Amsterdam: University of Amsterdam, 2015.

[13] Fox, R. A. and Jacewicz, E. "Cross-dialectal variation in formant dynamics of American English vowels," *The Journal of the Acoustical Society of America,* 126 (5), 2603-2618, 2009.

[14] Asu, E. L., Lippus, P., Niit, E. and Türk, H. "The Acoustic Characteristics of Monophthongs and Diphthongs in the Kihnu Variety of Estonian," *Linguistica Uralica*, XLVIII, 3, 2012.

[15] Mayr, R. and Davies, H. "A cross-dialect acoustic study of the monophthongs and diphthongs of Welsh," *Journal of the International Phonetic Association*, 41, 1-25, 2011.

[16] Mauder, E. and Van Heuven, V. "On the rise and fall of the Spanish diphthongs," In C. Cremers, & M. Den Dikken (Eds.), *Linguistics in the Netherlands*, p.171-182. Amsterdam: John Benjamins, 1996.

[17] Jaggers, Z. S. "Acoustic cues to the [j]-[i] distinction in American English," *Talk presented at the 170th Meeting of the Acoustical Society of America (ASA 170)*. Jacksonville, 2-6 November, 2015.