*Computer Assisted Language Learning Electronic Journal (CALL-EJ), 23*(1), 401-424, 2022

# English Diphthong Characteristics Produced by Thai EFL Learners: Individual Practice Using PRAAT

Narinthorn S. Behr (narinthorn@g.swu.ac.th)
Srinakharinwirot University, Bangkok, Thailand

## Abstract

Thai EFL learners may experience difficulty with the pronunciation of eight English diphthongs. It was considered that PRAAT, a speech analysis program, could assist learners to improve their performance. The purpose of the study was to investigate the pronunciation of these eight English diphthongs by a group of Thai EFL learners before and after one month of individual practice using PRAAT, after which participants would be asked to complete a feedback survey. F1 and F2 values as well as duration and rate of change (ROC) in diphthongs before and after training were measured and statistically compared with the *t*-test at $p < .05$. Results indicated that vowel duration was significantly longer after training. The first vocalic element appeared to be the longest part of the diphthong, and similar quality in this phase was observed before and after the training. Even though the learners seemed to struggle with the diphthongs /eɪ/, /əʊ/ and /eə/, the results for /eɪ/ after training exhibited values significantly greater than those for /əʊ/ and /eə/. Learners' feedback mostly reflected their satisfaction with PRAAT as a tool for pronunciation practice. It could be concluded that PRAAT offers practical benefits to Thai EFL learners.

*Keywords:* diphthongs, duration, formant frequency, PRAAT, Thai EFL learners

## Introduction

The dissimilarity between native and target languages may trigger difficulties during the language-learning process (Lightbown & Spada, 2006). Learners of English as a foreign language (EFL) may struggle with their performance (Kelly, 2000). Diphthong vowel articulation could be one possible area of difficulty. There are eight diphthong phonemes in Received Pronunciation (Roach, 2004), whereas only three diphthongs have been identified in the Thai vowel system (Tingsabadh & Abramson, 1993). Even though the number of vowel characteristics of the three Thai diphthongs appears to exceed the number of representations (Roengpitya, 2002), Thai EFL learners tend to equate some English diphthongs with monophthongs (Iadkert & Hashim, 2020; Tsukada, 2008).

For EFL learners, correct and clear pronunciation is important in language learning (Kriedler, 1989). They should practice accurate pronunciation in the target language to achieve unambiguous communication (Priya & Kumar, 2020). Problems experienced by the Thai EFL learners included difficulties pronouncing some consonants and monophthongs as well as the diphthongs /eɪ/, /au/ and /eə/ (Dee-in, 2006). Previous research has found that achieving correct diphthong vowel articulation was a common problem for Thai EFL learners due to differences between the English and Thai vowel systems (Sahatsathatsana, 2017). Thai EFL learners may find it easier to achieve clear pronunciation in English in cases where similar sounds exist in Thai, whereas the presence of unfamiliar sounds may hinder their L2 performance.

Technology can help EFL learners to achieve improved pronunciation. Previous studies have shown the positive impact of technology use on the performance of EFL learners (Brett, 2004; Fouz-González, 2020; Offerman & Olson 2016; Pardede, 2018; Wilson, 2008). Thai EFL learners were able to produce more accurate stress and intonation after four weeks of practice using e-learning (Yangklang, 2013). In a more recent study, the use of a mobile application offered by Google helped learners to pronounce some difficult words more accurately (Wongsuriya, 2020). A questionnaire and semi-structured interview from these studies on Thai EFL learners confirmed that technology could raise learners' level of engagement to their pronunciation practice. However, a review of related studies concerning research on technology in EFL education in Thailand published between 2004 and 2013 found that research on technology being used to improve speaking skills appeared with the lowest frequency (Deerajviset, 2014).

To enhance pronunciation practice, the use of technology should be more widely promoted in Thai EFL teaching and learning. PRAAT is a free software package that serves as a research tool for speech analysis and synthesis (Boersma & Weenink, 2009). This program can provide visual spectrograms as direct feedback to learners. It can offer benefits to EFL learners, helping them to improve their pronunciation (Aramipoor & Gorjian, 2018; Gorjiana, Hayati & Pourkhonic, 2013; Motohashi-Saigo & Hardison, 2009; Olson, 2014; Saito, 2007). In relation to the practical use of PRAAT with Thai EFL learners, Osatananda and Thinchan (2021) posited positive outcomes for learners on suprasegmental features. Even though PRAAT is well-known among phoneticians, it remains little-used in the teaching of pronunciation to Thai EFL learners. The purpose of the research project was to investigate the effects of the use of PRAAT on English pronunciation among a group of Thai EFL learners. In this study, participants had to practice their production of eight English diphthongs for one month using PRAAT. The diphthongs produced before and after this one-month practice period were acoustically analyzed in terms of first (F1) and second (F2) formant frequencies, duration, and rate of change (ROC), and statistically studied with the *t*-test at $p < .05$. General feedback on PRAAT was also solicited after training. The results may indicate prominent features of the diphthongs produced by the Thai EFL learners as well as students' satisfaction levels after the one-month practice period on PRAAT. It is hoped that the results will help to determine whether PRAAT technology offers a practical option for the implementation of pronunciation pedagogy for Thai EFL learners.

## Literature Review

### Diphthongs in English and Thai

One of the problems with Thai EFL pronunciation may stem from diphthong vowel articulation (Sahatsathatsana, 2017). A diphthong could be described as a combination of two individually pronounced vowels. The tongue moves from one position to another to produce a diphthong. Articulation can be displayed on a spectrogram and unquestionably reflected via the formant frequencies (Fry, 1979; Ladefoged, 2006). In Received Pronunciation (hereafter RP), which is extensively used in dictionaries and English textbooks published in Britain (Pennock-Speck, n.d.), there are eight diphthong representations (Roach, 2004). These could be described as five closing and three centering diphthongs. The closing diphthongs /eɪ/, /aɪ/, /ɔɪ/, /əʊ/, and /aʊ/ show movement from a more open quality in the first vocalic element towards a less open one in the second. The centering diphthongs /ɪə/, /eə/, /ʊə/ are articulated with the centering vowel /ə/ as their second element. On the contrary, there are three phonemic diphthongs, /ia/, /ɯa/ and /ua/, in the Thai language (Tingsabadh & Abramson, 1993). Roengpitya (2002) found that the first vocalic element showed the longest duration. The twenty-five percent of duration could indicate the prominent cue to differentiate the three diphthongs. The production and perception study revealed that the three phonemic diphthongs occurred with short and long phonetic properties (Roengpitya, 2002).

**Production of "New" and "Similar" phones among EFL learners**

The absence of some target language sounds in native languages may lead to mispronunciation among EFL learners. EFL learners may articulate similar sounds in their L1 and L2 performances instead of producing new ones. Flege (1987) hypothesized that EFL learners may face an upper limit with respect to their phonetic approximation. Even though they are capable of producing "new" phones, they may articulate L2 sounds using "similar" features from L1 phones. Trudgill and Hannah (1994) stated that Indian speakers tend to reduce the number of British English vowels in words, pronouncing some diphthongs /eɪ/ and /əʊ/ as the single vowels /e:/ and /o:/, as in their native languages. EFL learners with Chinese or Malay as their L1 showed a tendency to pronounce English diphthongs similarly to monophthongs in their native languages (Deterding, 2000). Sumbayak (2009) identified substitution of the English diphthongs /eɪ/ and /oʊ/ with the long vowels /e/ and /ɒ/ among EFL learners whose native language is Indonesian. A study of Pakistani EFL learners found the diphthongs /ʊə/ and /əʊ/ to be merged and articulated as the monophthong /o/, and the diphthong /eɪ/ to be produced as /e/ (Farooq & Mahmood, 2017). By contrast, familiarity with certain diphthongs is likely to help learners avoid mispronunciation. For example, in another study involving Indonesian EFL learners, it was found that the presence of the diphthong /ɔɪ/ in both Indonesian and English leads to fewer pronunciation mistakes involving this diphthong in English (Saadah & Ardi, 2020).

**Production of diphthongs among Thai EFL learners**

Thai EFL learners may employ existing sounds and features in the Thai language when speaking English. Tsukada (2008) analyzed the rate of change (ROC) of /eɪ/ and /oʊ/, comparing Australian speakers (hereafter AUS) and Thai EFL learners who had been living in Australia. ROC results clearly showed greater diphthong movement among members of the AUS group than was the case for the Thai EFL learners. The English diphthongs /eɪ/ and /oʊ/ were substituted with the long monophthong vowels /e:/ and /o:/ among the Thai learners. This paralleled the results found in Iadkert and Hashim (2020). The study showed that the diphthongs with the greatest change in F1 values were /ɔɪ/ and /ɪə/. The lower ROC of the three diphthongs /eɪ/, /əʊ/, /eə/ meant that they were likely to manifest as the monophthongs /eː/, /oː/ and /ɛː/ respectively. Contrastive vowel length in Thai may cause learners to produce durational exaggeration in their English performances, as pointed out in Tsukada (2009). When compared with native speakers, short vowels produced by Thai EFL learners had shorter durations, while long vowels had longer durations. The overuse of duration was also revealed in a study of English monophthongs produced by Thai EFL learners (Pillai, 2012).

**The use of Technology in pronunciation practice**

In line with a general expansion in the use of technology in education, computer-assisted language learning (CALL) can be effectively applied in the context of second language acquisition (SLA) (Chapelle, 2004). Technology and feedback elements seem to be important for language learners. Previous studies have demonstrated the positive impact of PRAAT on EFL learners. PRAAT is a free software program that is primarily used for speech analysis and synthesis (Boersma & Weenink, 2009). Learners can practice their pronunciation and observe their performances directly on the program. One study found that Japanese EFL learners produced English vowel sounds in an increasingly accurate manner after observing visual images. They were able to compare their performances with those of native speakers (Saito, 2007). PRAAT's spectrogram display helped the Japanese EFL learners to distinguish the differences between the consonants /ɾ/, /l/ and /r/, and to detect vowel length and voice onset time (VOT) in English (Wilson, 2008). In another study, Motohashi-Saigo and Hardison (2009) claimed that L2 Japanese beginners were more aware of durational differences after detecting the waveforms, while Olson (2014) explained the usefulness of the visual feedback paradigm when teaching consonants in L2 classrooms. A further study reached a similar finding, confirming significant improvement through the use of the visual feedback paradigm (VFD) in the transition from carrier sentences to spontaneous speech among Spanish L2 learners (Offerman & Olson, 2016). In another study, Li (2019) confirmed that using PRAAT in class benefited students. They were able to observe their performance, get feedback promptly and correct their mistakes.

**The use of technology for pronunciation practice among Thai EFL**

A study that reviewed the use of technology in Thai EFL classrooms published between 2004 and 2013 found that the highest frequencies of use were for vocabulary (23%), followed by reading (19.4%), grammar (15.8%), writing (13.3%), listening (9.1%), speaking (6.7%), and other integrated skills or areas (12.7%) (Deerajviset, 2014). Since that period, there has likely been an increase in the use of technology for pronunciation practice as well as an increase in positive outcomes in relation to learners' performance. As mentioned above, Yangklang (2013) reported an improvement in students' stress and intonation as a result of participation in e-learning activities. A pronunciation test indicated more accurate performance after the training, participants feeling satisfied with the content of the program. Wongsuriya (2020) posited the positive effects of using a Google translation app on Thai EFL learners. EFL learners in remote areas were able to improve their pronunciation and exhibited positive attitudes towards the use of technology. In another project, Arunsirot (2020) applied Augmented Reality (AR) technology to the practice of pronunciation. Participants in the experimental group were provided with support using AR technology when practicing consonant production. The follow-up test indicated a significant improvement among learners as a result of the use of AR technology. Learners also recorded the highest level of satisfaction in relation to the use of this technology. In a further study, Moxon (2021) examined Thai EFL learners' pronunciation on an online platform via SpeechAce. Participants were able to view their feedback and scores, which helped them to improve their pronunciation. Positive improvement was established in the treatment group. The existence of visual feedback was likely to encourage them to improve their performance. A further study by Osatananda and Thinchan (2021) demonstrated the positive impact of PRAAT use on Thai EFL learners. In the study, PRAAT training supplemented suprasegmental pronunciation training outside the classroom. Overall, participants exhibited positive feedback in their journal reflections, even though some may have encountered difficulties at the beginning of the program's operation.

Despite the benefits shown in the abovementioned research findings, PRAAT remains little-used in the context of pronunciation practice in Thai EFL classrooms. In line with a general expansion in the use of technology in Thai EFL teaching and learning, PRAAT may provide significant benefits to learners, helping them to improve their pronunciation, encouraging them to practice their performance, and triggering positive outcomes. It is hoped that the results of the research project discussed in this paper could lead to greater support for the use of technology in pronunciation pedagogy in Thailand.

## Methods

### Participants

Thirty Thai EFL undergraduate students in an English phonetics class, aged 19-22 and including twenty-three females and seven males[1] (see Table 1), volunteered to participate in the study.

Table 1
*Overview of participant information*

| Participant | n | % |
|---|---|---|
| Gender | | |
| Female | 23 | 76.67 |
| Male | 7 | 23.33 |
| Age (19-22) | | |
| 19 | 21 | 70 |
| 20 | 6 | 20 |
| 21 | 1 | 3.33 |
| 22 | 2 | 6.67 |
| Education | | |
| University level | 30 | 100 |

[1] No statistical differences were observed in the results for female and male participants. Therefore, the combined results were used in further analysis.

## Word list

Word lists were provided and used for two purposes: a) to assist the researcher with acoustic analysis: and b) to assist the students during their practice sessions. The lists were created from the eight diphthong representations /eɪ/, /aɪ/, /ɔɪ/, /aʊ/, /əʊ/, /ɪə/, /eə/ and /ʊə/, and included open (CV) and closed (CVC) syllables. Most words were monosyllabic, with a few bisyllabic words containing the diphthong in the stressed syllable. Syllables with a final /r/ were classified as open syllables, due to the absence of /r/ in British English. For acoustic analysis, ten words were included for each diphthong. A further ten words were added to the individual practice list. Therefore, there was a total of eighty test words for the acoustic study and 160 for the individual practice component.

## Data Analysis

### *Acoustic Study*

## Recording

The participants recorded the eighty words on the list three times randomly before and after the one-month period of individual practice on PRAAT with a 22,500-sampling rate. In total, each student practiced 7,200 words before the training and 7,200 after the training.

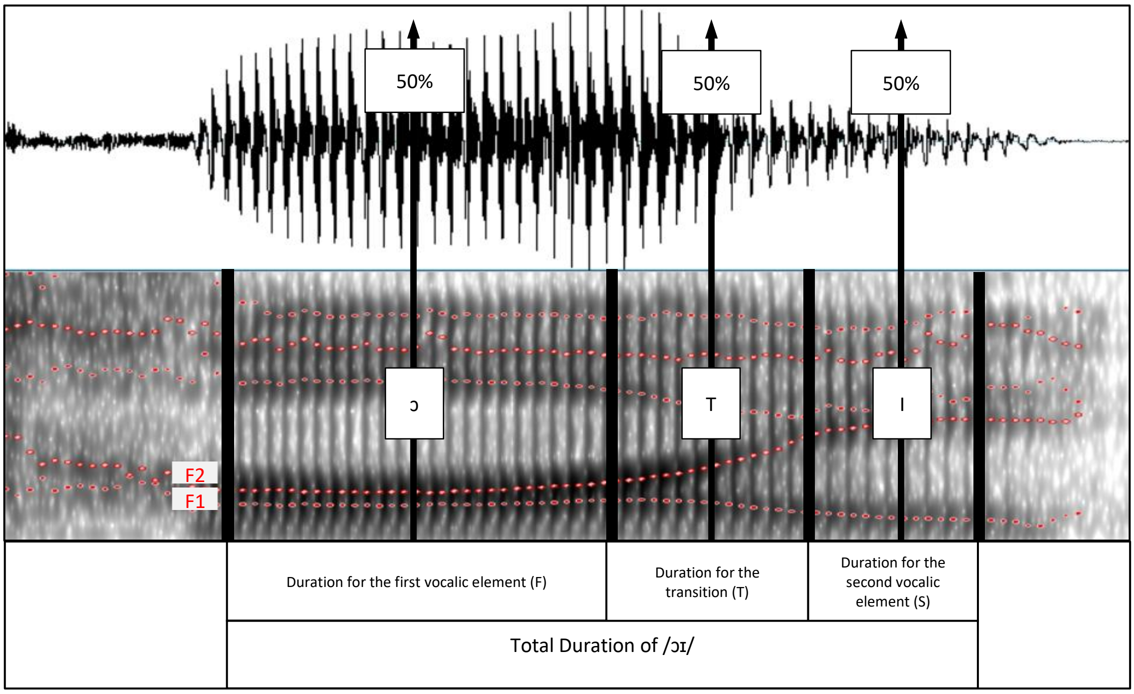
## Acoustic Analysis

### *Acoustic measurement*

Vowels were acoustically analyzed using PRAAT Version 6.1.07 (Boersma & Weenink, 2019). PRAAT was developed by Paul Boersma and David Weenink. It is a free software package that enables the recording, editing, and analysis of sounds in terms of their acoustic values, such as formant frequency, fundamental frequency, and intensity. After opening a sound on PRAAT, the user is required to annotate, then click on TextGrid to divide the tier to obtain the acoustic values. Both sound and TextGrid objects need to be selected before the file can be edited. The user has to label the waveform and spectrogram in addition to setting the particular acoustic values as “Show”. The values in a specific interval will then be listed. The user can select the total values in the interval or any points in time to be saved in a text file and analyzed in an Excel program. PRAAT also offers users the option of writing a script to help extract the values automatically. However, this was not performed in this study.

Visual identification and listening judgment were used to avoid the influence of initial and final consonants on vowels. Diphthongs displayed with one steady state were calculated separately and treated as monophthongs. Due to noise problems, four hundred words were omitted.

Figure 1
*Example of the acoustic measurement of diphthongs*




Each diphthong was divided into three parts: (1) the first vocalic element or the first vowel phase (F); (2) a transition (T); and (3) the second vocalic element or the second vowel phase (S) (see Figure 1). The acoustic parameters were studied as follows:

1) 50% of the first formant (F1) and the second formant (F2) frequency from each interval were acoustically measured in Hertz (Hz).

2) The vowel duration in each phase and the total duration were measured in milliseconds (ms).

3) The rate of change (ROC) was analyzed. In this study, this was the difference in formant frequency values between two-time points (50% of the first vocalic element and 50% of the second one) which was then divided by the duration.

***Statistical study***

Each acoustic measurement was statistically analyzed with the *t*-test at $p < .05$ in terms of the values before and after training and between each phase as: (1) the first vocalic element vs. the transition (F:T); (2) the transition vs. the second vocalic element (T:S); and (3) the first vocalic vs. the second vocalic element (F:S).

***Questionnaire***

Participants completed a questionnaire after the training period to provide them with an opportunity to reflect on their use of PRAAT and provide feedback. The questionnaire included six questions. Five of the items featured a five-point Likert scale, participants being asked to provide responses ranging from “Strongly disagree = 1” to

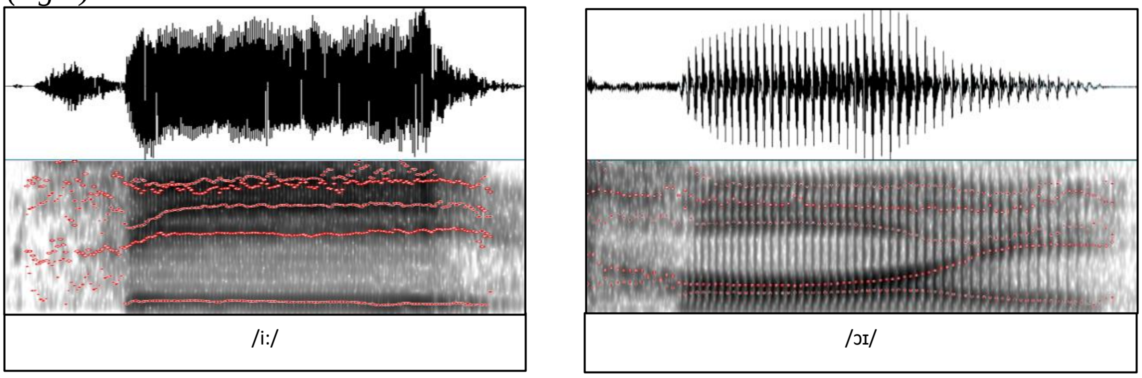
"Strongly agree = 5." The sixth item featured an open-ended question on the use of PRAAT for pronunciation practice.

## Training

### *During the first three weeks*

Participants studied basic knowledge related to articulatory and acoustic phonetics. Afterward, they learned how to operate the PRAAT program and display the visual program on a screen. They were instructed on how to observe formant movement in a diphthong, which differs from the stability of a formant in a monophthong (See Figure 2).

Figure 2
*Example of a spectrogram showing a monophthong /i:/ (left side) and diphthong /ɔɪ/ (right).*




### *One month after the in-class training*

Participants were required to practice their pronunciation for acoustic analysis on PRAAT, pronouncing the 160 words as individual practice every other day. They performed these activities outside the classroom at their own pace. The participants had to provide an update concerning their weekly progress during their phonetics class each week.

The diphthong characteristics produced by the Thai EFL learners before and after one month of self-practice on PRAAT were acoustically investigated, and the overall feedback was conducted after training.

# Results

The acoustic characteristics of the eight diphthong representations /eɪ/, /aɪ/, /ɔɪ/, /aʊ/, /əʊ/ /ɪə/, /eə/, and /ʊə/ produced by the Thai EFL learners before and after the period of training will now be examined in terms of F1 and F2 values, duration and ROC. Overall feedback on the use of PRAAT will then be presented.

## Acoustic results

***Formant frequency***

The *t*-test at $p < .05$ of F1 and F2 values shows overall differences. Differences in F1 and F2 values for the transition and second vocalic element of diphthongs before and after training are mostly significant. Values are relatively similar for the first phase of many diphthongs, except in the case of /eɪ/ (see Table 2). The F1 value for /eɪ/ is 528.21 Hz before training and 602.90 Hz after training (t (1723) = -4.33, $p = 0$) and its F2 values are 2175.42 Hz and 2237.38 Hz (t (1626) = -3.01, $p = 0$). The lack of stability in F1 and F2 at the beginning of /eɪ/, indicating a significant difference, may result from the smaller number of diphthongs in Thai. There are only three phonemic diphthongs /ia/, /ɯa/ and /ua/) in Thai (Tingsabadh & Abramson, 1993), whereas there are eight diphthong representations in English: /eɪ/, /aɪ/, /ɔɪ/, /əʊ/, /aʊ/, /ɪə/, /eə/, and /ʊə/ (Roach, 2004). The absence of /eɪ/ in Thai may have led to this result.

Table 2
*Mean values of F1 and F2 values at 50% in the first vocalic element (*indicates significant values)*

| Vowel | First vocalic element (F) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | F1 (Hz) | | | | F2 (Hz) | | | |
| | before | after | t | p | before | after | t | p |
| /eɪ/ | 582.21 | 602.90 | -4.33 | 0* | 2175.42 | 2237.38 | -3.01 | 0* |
| /aɪ/ | 943.18 | 942.89 | 0.04 | 0.48 | 1533.84 | 1544.92 | -1.11 | 0.13 |
| /ɔɪ/ | 692.97 | 701.04 | -1.29 | 0.10 | 1083.78 | 1065.17 | 1.82 | 0.03* |
| /aʊ/ | 927.99 | 942.52 | -1.92 | 0.03* | 1489.81 | 1494.66 | -0.46 | 0.32 |
| /əʊ/ | 650.64 | 642.27 | 1.44 | 0.07 | 1207.07 | 1229.58 | -1.78 | 0.04* |
| /ɪə/ | 411.57 | 413.14 | -0.39 | 0.35 | 2306.99 | 2331.94 | -0.97 | 0.17 |
| /eə/ | 667.54 | 670.98 | -0.51 | 0.30 | 1994.42 | 2016.97 | -0.90 | 0.18 |
| /ʊə/ | 462.31 | 474.01 | -2.41 | 0.01* | 1005.93 | 1006.78 | -0.04 | 0.48 |

Table 3
*Mean values of F1 and F2 values at 50% in the transition phase*

| Vowel | Transition (T) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | F1 (Hz) | | | | F2 (Hz) | | | |
| | before | after | t | p | before | after | t | p |
| /eɪ/ | 512.35 | 493.16 | 3.86 | 0* | 2211.71 | 2309.79 | -3.70 | 0* |
| /aɪ/ | 512.35 | 493.16 | 3.86 | 0* | 2211.89 | 2310.02 | -3.70 | 0* |
| /ɔɪ/ | 701.44 | 667.41 | 5.77 | 0* | 1476.40 | 1577.52 | -5.85 | 0* |
| /aʊ/ | 819.43 | 801.39 | 2.04 | 0.02* | 1298.91 | 1300.93 | -0.18 | 0.43 |
| /əʊ/ | 575.43 | 562.52 | 2.54 | 0.01* | 1038.57 | 1082.60 | -4.24 | 0* |
| /ɪə/ | 544.35 | 546.34 | -0.39 | 0.35 | 2075.75 | 2048.64 | 1.36 | 0.09 |
| /eə/ | 700.69 | 683.41 | 2.81 | 0* | 1989.62 | 1933.10 | 2.74 | 0* |
| /ʊə/ | 539.91 | 538.95 | 0.21 | 0.42 | 1063.25 | 1028.27 | 2.45 | 0.01* |

Table 4
*Mean values of F1 and F2 values at 50% in the second vocalic element*

| Vowel | Second vocalic element (S) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | F1 (Hz) | | | | F2 (Hz) | | | |
| | before | after | t | p | before | after | t | p |
| /eɪ/ | 451.70 | 400.25 | 10.33 | 0* | 2428.02 | 2531.30 | -5.00 | 0* |
| /aɪ/ | 558.24 | 528.16 | 3.06 | 0* | 2179.47 | 2202.39 | -1.03 | 0.15 |
| /ɔɪ/ | 507.99 | 483.79 | 3.67 | 0* | 2102.32 | 2234.99 | -5.88 | 0* |
| /aʊ/ | 570.94 | 560.89 | 1.30 | 0.10 | 1118.42 | 1092.11 | 1.94 | 0.03* |
| /əʊ/ | 505.65 | 476.46 | 5.85 | 0* | 946.80 | 1056.28 | -9.24 | 0* |
| /ɪə/ | 694.72 | 678.34 | 2.27 | 0.01* | 1760.86 | 1685.99 | 4.92 | 0* |
| /eə/ | 698.95 | 661.37 | 5.57 | 0* | 1917.05 | 1806.89 | 5.75 | 0* |
| /ʊə/ | 614.61 | 603.86 | 1.70 | 0.04* | 1191.89 | 1231.04 | -2.83 | 0* |

Formant frequencies in the transition phase before and after training exhibit significant differences for most vowels (see Table 3). However, /ɪə/ demonstrates non-significant F1 and F2 values. F1 values for /ɪə/ before and after training are 544.35 Hz and 546.36 Hz respectively (t (1614) = -0.39, *p* = 0.35). Its F2 values are 2075.75 Hz and 2048.64 Hz (t (1614) = 1.36, *p* = 0.09).

Differences in values for most vowels can be observed for the second vocalic element before and after training (see Table 4). Nonetheless, F1 values for /aʊ/ and F2 values for /aɪ/ show no significant differences at *p* < .05. F1 values for /aʊ/ before and after training are 570.94 Hz. and 560.89 Hz respectively (t (1798) = 1.30, *p* = 0.10). F2 values for /aɪ/ are 2179.47 Hz and 2202.39 Hz (t (1798) = -1.03, *p* = 0.15). The presence of /a/+/w/ and /a/+/j/ in Thai may have contributed to the stability of vowel articulation observed in the production of the English /aʊ/ and /aɪ/ before and after training.

Figure 3
*The vowel area of diphthongs produced by the Thai EFL learners before training*

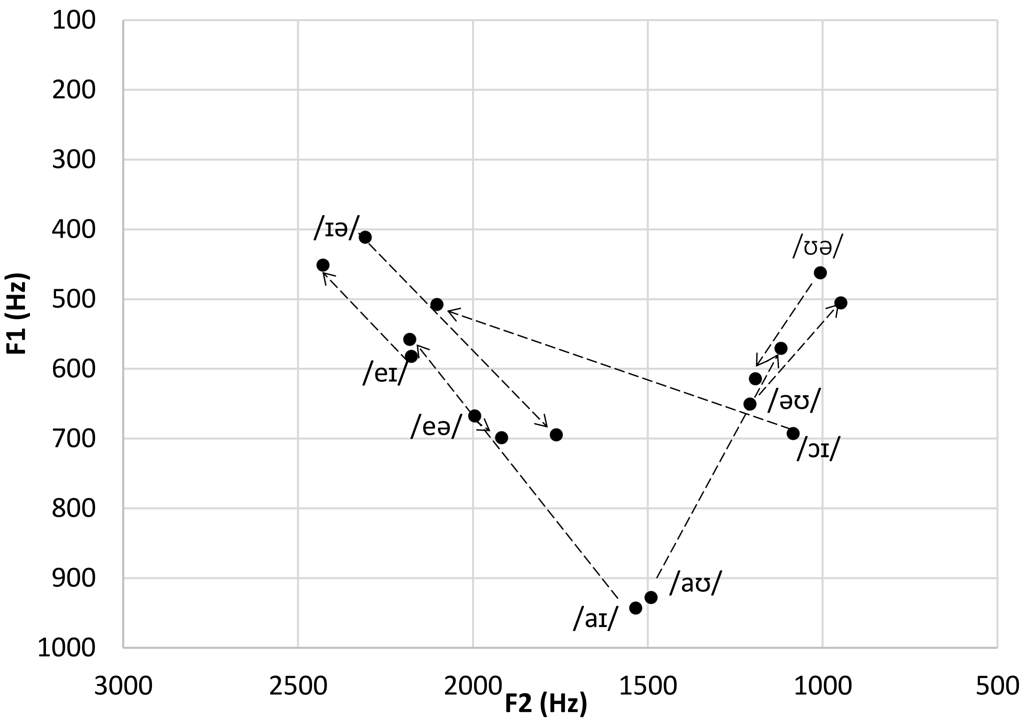
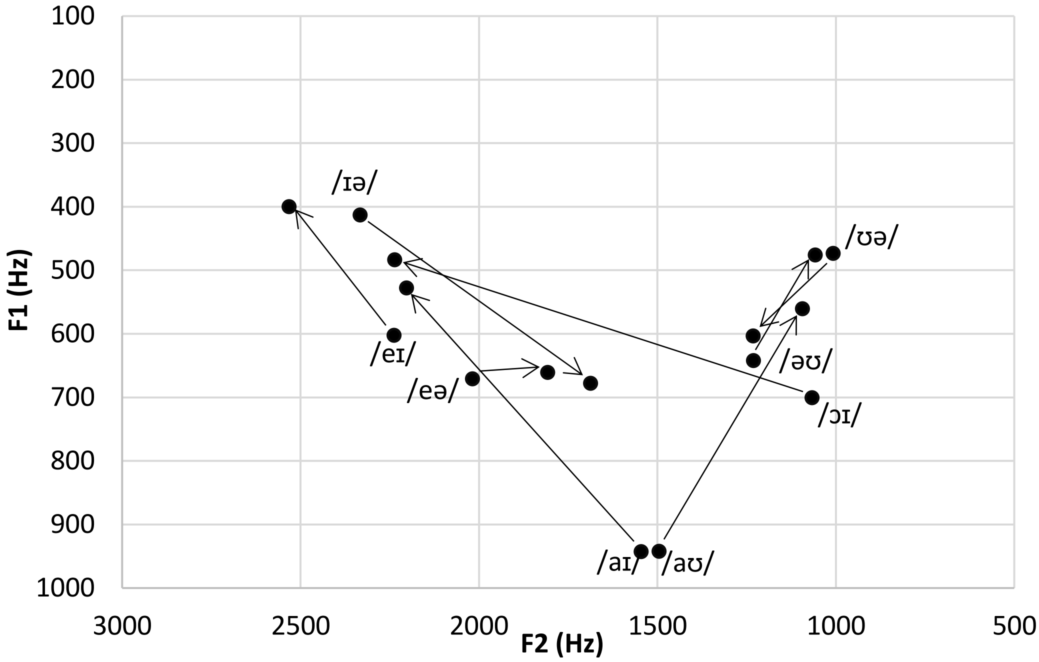


Figure 4

*The vowel area of diphthongs produced by the Thai EFL learners after training*




Even though differences in the F1 and F2 values for many diphthongs before and after training are significant for the transition and second vocalic element, they display relatively similar patterns in the vowel area, except in the cases of /eə/ and /əʊ/ (see Figure 3 and Figure 4). Production of /eə/ tends to have a more open quality before training but moves more to the back and center by the end of training. For /əʊ/, formant frequencies in the second element lead to greater space between the production of first and second vowels after training. The formant values for /ʊ/ in /əʊ/ seem to be similar to those for /ʊ/ in /ʊə/ after training.

The monophthongs /e:/, /o:/, /ɛ:/, which sometimes replace /eɪ/, /əʊ/ and /eə/, were acoustically analyzed. A comparison of F1 and F2 values for each phase for the diphthongs /eɪ/, /əʊ/ and /eə/, and for 50% of the monophthongs /e:/, /ɛ:/ and /o:/ indicates significant differences (see Table 5). Differences in values for /eɪ/ and /e:/ before and after training are all significant, whereas some similar values can be observed in the cases of /əʊ/ to /o:/ and /eə/ to /ɛ:/. Differences in F2 values for the second vowel phase in the case of /eə/ and /ɛ:/ are not significant (t (876) = -0.66, $p$ = 0.51) before training, but are significant after training (t (624) = -8.35, $p$ = 0). Differences in F1 values between /əʊ/ and /o:/ for the transition phase are not significant (t (15) = 1.93, $p$ = 0.07), as is the case for differences in F2 values between /eə/ and /ɛ:/ (t (779) = 1.17, $p$ = 0.24) for the first vowel phase after training. These findings may reflect similar articulation of these vowels in terms of tongue height for /əʊ/ and /o:/ and in terms of tongue advancement in /eə/ and /ɛ:/ after training.

Table 5

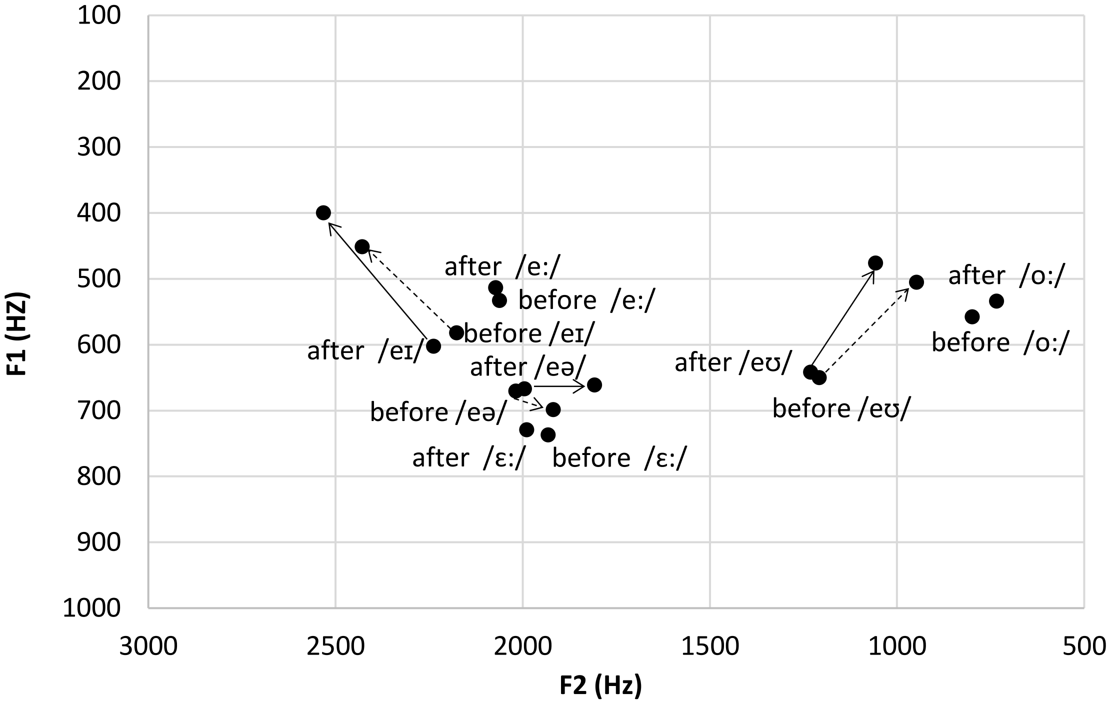
*F1 and F2 values for each phase of diphthongs and for 50% of replaced monophthongs*

| Value | Vowel | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | /eɪ/ | /eː/ | t | p | /əʊ/ | /oː/ | t | p | /eə/ | /ɛː/ | t | p |
| First vocalic element (F) | | | | | | | | | | | | |
| Before | | | | | | | | | | | | |
| F1 | 582.21 | 532.88 | 7.61 | 0* | 650.64 | 558.20 | 13.54 | 0* | 667.54 | 737.43 | -10.77 | 0* |
| F2 | 2175.42 | 2061.30 | 5.86 | 0* | 1207.07 | 798.30 | 27.47 | 0* | 1994.42 | 1931.32 | 2.76 | 0* |
| After | | | | | | | | | | | | |
| F1 | 602.90 | 514.10 | 19.31 | 0* | 642.27 | 534.34 | 12.66 | 0* | 670.98 | 729.74 | -7.74 | 0* |
| F2 | 2237.38 | 2071.16 | 11.56 | 0* | 1229.58 | 734.82 | 25.07 | 0* | 2016.97 | 1988.56 | 1.17 | 0.24 |
| Transition (T) | | | | | | | | | | | | |
| Before | | | | | | | | | | | | |
| F1 | 512.35 | 532.88 | -3.20 | 0* | 575.43 | 558.20 | 2.84 | 0* | 700.69 | 737.43 | -6.28 | 0* |
| F2 | 2211.71 | 2061.30 | 6.66 | 0* | 1038.57 | 798.30 | 17.55 | 0* | 1989.62 | 1931.32 | 3.01 | 0* |
| After | | | | | | | | | | | | |
| F1 | 493.16 | 514.10 | -4.30 | 0* | 562.52 | 534.34 | 1.93 | 0.07 | 683.41 | 729.74 | -6.07 | 0* |
| F2 | 2309.79 | 2071.16 | 12.72 | 0* | 1082.60 | 734.82 | 18.30 | 0* | 1933.10 | 1988.56 | -2.36 | 0.02* |
| Second vocalic element (S) | | | | | | | | | | | | |
| Before | | | | | | | | | | | | |
| F1 | 451.70 | 532.88 | -12.50 | 0* | 505.65 | 558.20 | -8.20 | 0* | 698.95 | 737.43 | -6.09 | 0* |
| F2 | 2428.02 | 2061.30 | 19.45 | 0* | 946.80 | 798.30 | 11.25 | 0* | 1917.05 | 1931.32 | -0.66 | 0.51 |
| After | | | | | | | | | | | | |
| F1 | 400.25 | 514.10 | -24.10 | 0* | 476.46 | 534.34 | -6.10 | 0* | 661.37 | 729.74 | -8.81 | 0* |
| F2 | 2531.30 | 2071.16 | 30.37 | 0* | 1056.28 | 734.82 | 16.04 | 0* | 1806.89 | 1988.56 | -8.35 | 0* |

F1 and F2 values indicate a relative closeness in quality to one of the two-vowel combinations in the diphthong vowel area (see Figure 5). This could help explain the frequent mispronunciation of these vowels especially /eə/ to /ɛ:/ before training. Even though the diphthongs /eɪ/ and /əʊ/ seem to equate to /e:/ and /o:/, their values after training are likely to be separate from those of the monophthongs. /eə/ could be the most difficult one to differentiate from /ɛ:/.

Figure 5

*Differences in F1 and F2 values between the diphthongs /eɪ/, /əʊ/ and /eə/ and the monophthongs /e:/, /ɛ:/ and /o:/*

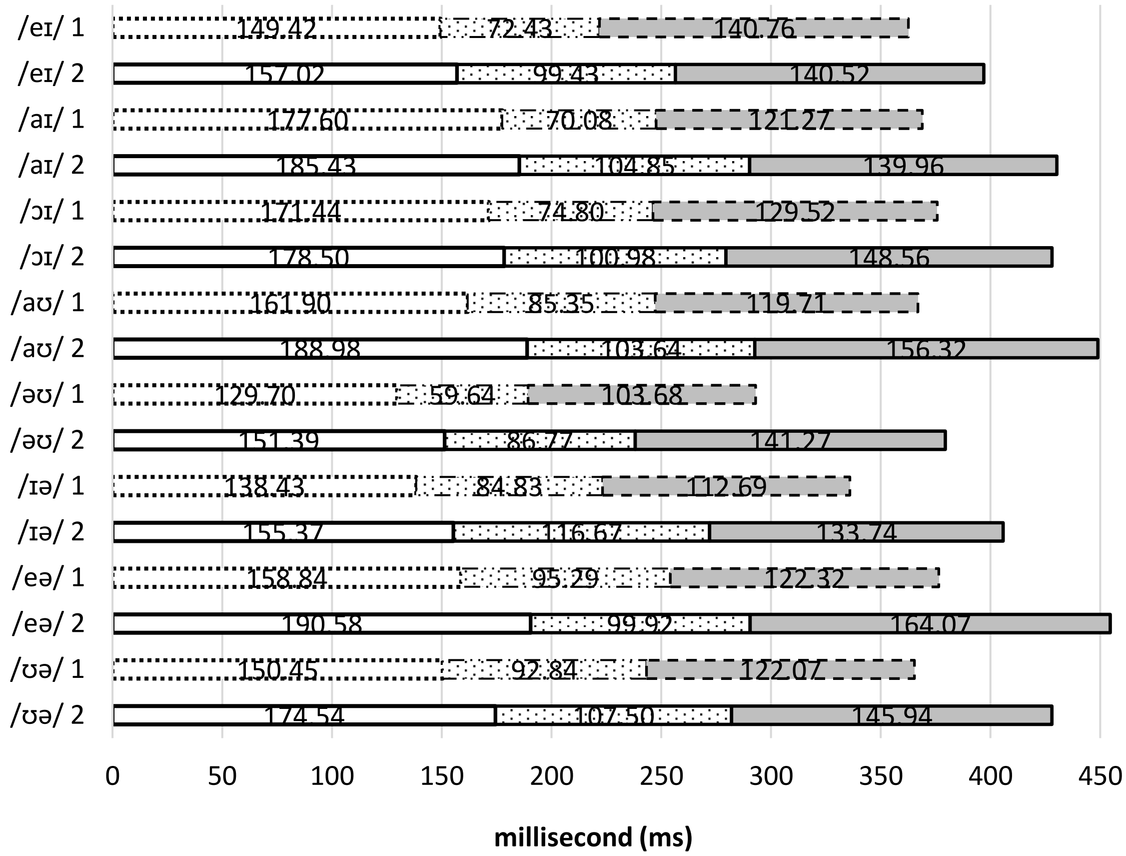



### *Duration*

The total duration of the diphthongs /eɪ/, /aɪ/, /ɔɪ/, /aʊ/, /əʊ/, /eə/, /ɪə/, and /ʊə/ after training is significantly longer than before training (see Figure 6). The diphthong /eə/ shows the longest duration, both before and after training, with figures of 376.45 ms and 454.57 ms respectively (t (1125) = -7.70, $p$ = 0). The diphthong /əʊ/ has the shortest duration, registering 293.02 ms before and 379.43 ms after training (t (1664) = -8.47, $p$ = 0). The duration of /eɪ/ is 362.61 ms before and 396.97 ms after training (t (952) = -1.90, $p$ = 0.03). For /aɪ/, the duration is 368.95 ms before and 430.24 ms after training (t (1796) = -10.33, $p$ = 0), for /ɔɪ/ 375.76 ms before and 428.04 ms after training (t (1796) = -7.86, $p$ = 0), and for /aʊ/ 366.96 ms before and 448.94 ms after training (t (1796) = -10.98, $p$ = 0). The duration of /ɪə/ is 335.95 ms before and 405.78 ms after training (t (1616) = -8.50, $p$ = 0), and for /ʊə/ 365.36 ms before and 427.98 ms after training (t (1616) = -11.17, $p$ = 0).

Figure 6

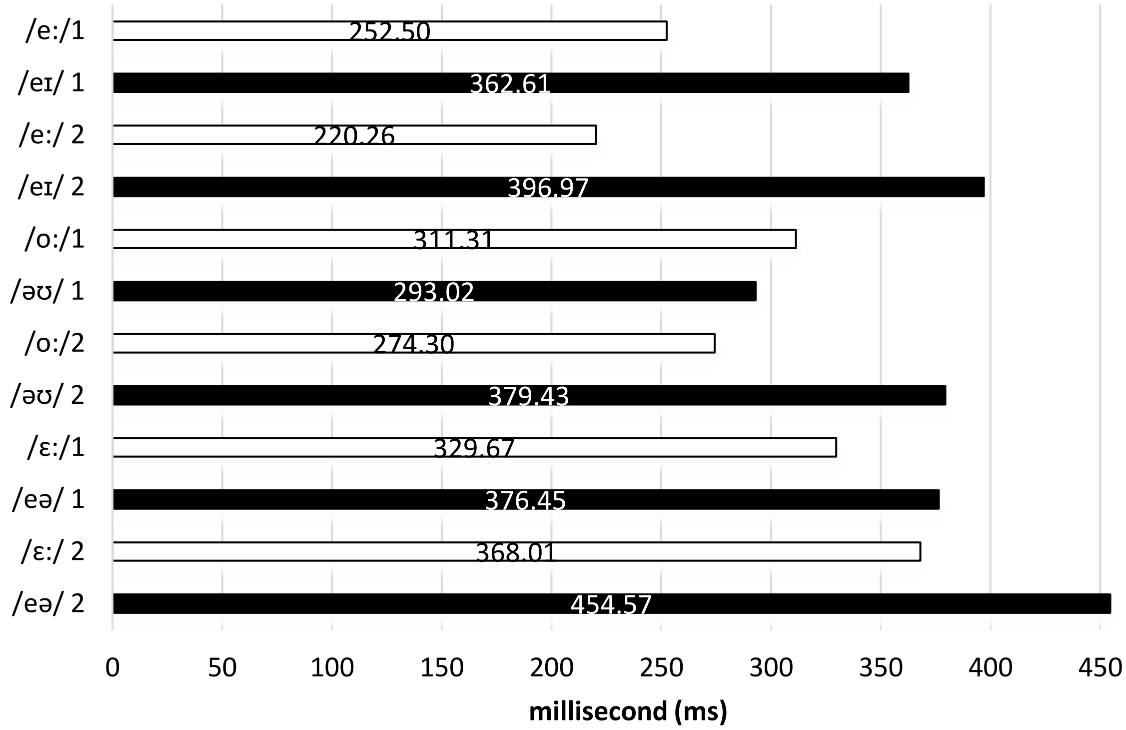
*Duration of diphthongs in each phase before and after training (1 indicates values before training and 2 indicates values after training)*




The participants sometimes mispronounced the diphthongs /eɪ/, /əʊ/ and /eə/ as the monophthongs /e:/, /o:/, and /ɛ:/. Therefore, the durations of these monophthongs were analyzed separately and compared with /eɪ/, /əʊ/ and /eə/. The results reveal that the durations of /e:/ and /o:/ are significantly shorter after training, whereas /ɛ:/ is significantly longer (see Figure 7). The vowel /e:/ lasts 252.50 ms before training and 220.26 ms after training (t (79) = 2.49, $p$ = 0.01), /o:/ for 311.31 ms before training and 274.30 ms after training (t (128) = 2.63, $p$ = 0), and /ɛ:/ for 329.67 ms before training and 368.01 ms after training (t (610) = -6.97, $p$ = 0). After training, the vowel /e:/ displays the shortest duration.

Figure 7

*The duration between the diphthongs /eɪ/, /əʊ/ and /eə/ and the monophthongs /e:/, /o:/ and /ɛ:/ before and after training (1 indicates values before training and 2 indicates values after training)*




Some diphthongs show similar or shorter durations before training, but all are significantly longer than the monophthongs after training. The durations of /eɪ/ and /e:/ are 362.61 ms and 252.50 ms respectively before training (t (468) = 5.27, p=0) and 396.97 ms and 220.26 ms /eɪ/and /e:/ (t (54) = 23.33, p=0) respectively after training. For /eə/ and /ɛ:/, the durations are 376.45 ms and 329.67 ms respectively before training (t (755) = -7.70, p=0) and 454.57 ms and 368.01 ms (t (873) = 10.36, p=0) respectively after training. These two diphthongs are significantly longer than the monophthongs before and after training. However, /əʊ/ is shorter than /o:/ before training, with similar respective lengths of 293.02 ms and 311.31 ms (t (175) = -1.25, $p$ = 0.11). After training, /əʊ/ lasts longer than the monophthong /o:/, with respective durations of 379.43 ms and 274.30 ms (t (153) = 11.24, p=0). In addition, a significant difference in duration between /e:/ and /eɪ/ can be observed after training. It can be hypothesized that the participants emphasized this feature to produce the diphthongs and monophthongs distinctively.

A statistical comparison was conducted in relation to the F1 and F2 values for each phase as well as the durations between phases: the first vocalic element vs the transition (F:T); the transition vs the second vocalic element (T:S); and the first vocalic element vs the second vocalic element (T:S). It was considered that such a comparison could help to clarify the characteristics of two-vowel combinations in diphthongs. The results show significant differences between phases for most vowels, except in the cases of /eɪ/, /əʊ/ and /eə/. Before training, /eɪ/ shows non-significant differences in F2 values for F:T (t (1736) = -1.80, $p$ = 0.07), and in duration for F:S (t (1678) = 0.51, $p$ = 0.61), but shows significant differences after training. Non-significant differences in F2 values for /əʊ/ for T:S (t (1750) = 1.83, *p* = 0.06), in duration for F:S (t (1582) = 1.08, *p* = 0.27), and in F1 values for /eə/ for F:T (t (1198) = -1.79, *p* = 0.07) and F:S (t (1198) = 1.34, *p* = 0.18) can be observed after training. For /eə/, non-significant differences in F1 values for F:S before and after training can also be observed.

***Rate of Change (ROC)***

A comparison of the ROC of F1 and F2 values before and after training revealed similar patterns of movement for most diphthongs (see Table 6 and Table 7). Some diphthongs (/eə/, /aɪ/ and /ɔɪ/) exhibit significant ROC values. For /eə/, the ROC of F1 values indicates is greater before training (160.35 Hz) than after training (43.70 Hz) (t (1117) = 3.61, *p* = 0). The ROC of F2 values for /ɔɪ/ is 4391.39 Hz before and 4625.43 Hz after training (t (1798) = -1.98, *p* = 0.02). The vowel /aɪ/ shows a statistical difference in the ROC of F1 values before and after training (-1747.26 Hz and -1620.47 Hz (t (1798) = -2.30, *p* = 0.02) respectively) and in the ROC of F2 values before and after training (2574.78 Hz and 2352.34 Hz (t (1798) = 1.90, *p* = 0.02) respectively).
Some vowels seem to vary in terms of the ROC of their F1 and F2 values, but such differences are not significant. The ROC of F1 values for /eɪ/ is -627.03 Hz before and -848.84 Hz after training (t (1648) = 8.14, *p* = 3.56), and for /ʊə/ 1226.7 Hz before and 552.55 Hz after training (t (753) = 11.43, *p* = 2.54). The ROC of F2 values for /əʊ/ is -1466.48 Hz before and -952.31 Hz after training (t (1631) = -7.25, *p* = 3.26), for /ɪə/ -2621.54 Hz before and -3339.60 Hz after training (t (1618) = 1.16, *p* = 0.12), for /eə/ -415.39 Hz before and -993.47 Hz after training (t (1094) = 4.63, *p* = 2.04), and for /ʊə/ 1385.72 Hz before and 907.76 Hz after training (t (1060) = 4.47, *p* = 4.34).

Table 6
*Rate of change (ROC) values for the F1 and F2 of closing diphthongs before and after training*

| Vowel | Before | After | t | p |
|---|---|---|---|---|
| /eɪ/ | | | | |
| F1 | -627.03 | -848.84 | 8.14 | 3.56 |
| F2 | 1196.93 | 1296.68 | -0.70 | 1.64 |
| /aɪ/ | | | | |
| F1 | -1747.26 | -1620.47 | -2.30 | 0.02* |
| F2 | 2574.78 | 2352.34 | 1.90 | 0.02* |
| /ɔɪ/ | | | | |
| F1 | -858.26 | -833.41 | -0.50 | 0.29 |
| F2 | 4391.39 | 4625.43 | -1.98 | 0.02* |
| /əʊ/ | | | | |
| F1 | -751.67 | -736.51 | -0.49 | 0.31 |
| F2 | -1466.48 | -952.31 | -7.25 | 3.26 |
| /aʊ/ | | | | |
| F1 | -1580.31 | -1527.14 | -1.18 | 0.12 |
| F2 | -1695.16 | -1591.75 | -1.60 | 0.05 |

Table 7

*Rate of change (ROC) values for the F1 and F2 of centering diphthongs before and after training*

| Vowel | Before | After | t | p |
|---|---|---|---|---|
| /ɪə/ | | | | |
| F1 | 1450.42 | 1446.09 | 0.02 | 0.49 |
| F2 | -2621.54 | -3339.60 | 1.16 | 0.12 |
| /eə/ | | | | |
| F1 | 160.35 | 43.70 | 3.61 | 0* |
| F2 | -415.39 | -993.47 | 4.63 | 2.04 |
| /ʊə/ | | | | |
| F1 | 1226.7 | 552.55 | 11.43 | 2.54 |
| F2 | 1385.72 | 907.76 | 4.47 | 4.34 |

**The questionnaire**

In their questionnaire responses (see Table 8), most participants (80%) agreed on the usefulness of PRAAT for practicing the pronunciation of diphthongs. The visual spectrogram seemed to assist 76.7% of participants in indicating the two-vowel movement of diphthongs. 73.3% of participants responded that they had used PRAAT to have a more interesting individual practice experience. Operating PRAAT seemed uncomplicated for more than 60% of participants. Almost 50% of the participants responded that they desired to continue individual practice using PRAAT, while 36.7% said that they may or may not continue to use the program. However, some participants (33.3%) thought that PRAAT was not simple to operate.

Table 8

*Learners' feedback on the use of PRAAT for pronunciation practice*

| Question | 5-point scale | | | | |
|---|---|---|---|---|---|
| | Strongly disagree | Disagree | Neutral | Agree | Strongly agree |
| PRAAT is useful software for diphthong practice. | 1 (3.3%) | 0 | 5 (16.7%) | 4 (13.3%) | 20 (66.7%) |
| PRAAT can increase your self-practice to be more interesting. | 0 | 2 (6.7%) | 6 (20%) | 6 (20%) | 16 (53.3%) |
| The visual spectrogram on PRAAT could help clarify the diphthong movement. | 0 | 1 (3.3%) | 6 (20%) | 9 (30%) | 14 (46.7%) |
| PRAAT is easy for you to operate. | 0 | 1 (3.3%) | 9 (30%) | 9 (30%) | 11 (36.7%) |
| If you have more time, you would like to continue your self-practice using PRAAT. | 1 (3.3%) | 4 (13.3%) | 11 (36.7%) | 8 (26.7%) | 6 (20%) |

For the open-ended question, responses tended to indicate that students found PRAAT to be a useful program. It helped them produce clearer pronunciation. The following examples illustrate their positive attitude to the program: "It is useful to learn how to pronounce with phonetic features," "Using PRAAT helped me to pronounce sounds more accurately," and "My pronunciation is better. The visual spectrogram makes it easier". Most participants mentioned the practical usefulness of PRAAT. The visual program certainly helped them to notice differences between two-vowel combinations in diphthongs. The program was easy to operate and also helped them to understand the physical properties of sounds. A few participants expressed the opinion that using PRAAT may not be the most suitable way for them to practice their pronunciation naturally. It is a little complicated. As one mentioned, "In my opinion, it helps us to focus on pronunciation more, but not in a natural way". Another wrote: "It is a really good program, but it is sometimes hard to use".

## Discussion

### The main feature of English diphthongs produced by Thai EFL learners

The results of this study exhibit that vowel duration among the group of Thai EFL learners is significantly longer after the period of PRAAT training. Vowel length distinction in the Thai language may have affected Thai EFL learners' production. This may explain why learners used this feature in assessing their performance after training. Duration in this study also seems to be overused, an assessment that can also be made in relation to previous studies on Thai EFL learners (Pillai, 2012; Tsukada, 2009). However, the overuse of duration tends to be a feature common to all syllable types in this study. In addition, the fact that the longest duration can be observed in the first vocalic element tends to confirm that Thai EFL learners emphasize the first part of a diphthong in their language performance. This phase is also indicated as an important cue for distinguishing the three phonemic diphthongs in Thai (Roengpitya, 2002). The significantly longer durations for every diphthong after training, possibly indicate that this is the main feature for Thai EFL learners.

For F1and F2 values as well as ROC, no abrupt changes can be observed for the Thai EFL learners after their period of individual practice on PRAAT. The results indicate less formant movement in diphthongs as shown in previous studies (Iadkert & Hashim, 2020; Tsukada, 2008). It is possibly more complicated to articulate different vowel combinations in terms of tongue height and advancement in the case of some diphthongs.

### Some potentially challenging diphthongs for Thai EFL learners

Familiarity with this type of two-sound combination from their native language may have helped the Thai EFL learners when producing certain sounds. In the results of the study, some diphthong representations do not appear to cause difficulty for Thai EFL learners. The combinations /ɪə/ and /ʊə/ in English may not be different from /ɪa/ and /ʊa/ in the Thai vowel system. It is possible that the learners produced their English /aɪ/, /aʊ/ and /ɔɪ/ by pronouncing the Thai monophthongs /a/ and /ɔ/ and adding a final /-j/ or /-w/.

On the contrary, the Thai EFL learners seemed to struggle with some diphthongs which do not exist in the Thai sound system. The presence of unfamiliar sounds was likely to trigger some obstacles in diphthong pronunciation. The diphthongs /eɪ/, /əʊ/ and /eə/ tend to be mispronounced as /e:/, /o:/ and /ɛ:/ respectively. Previous research has shown that some diphthongs are likely to become single vowels among EFL learners (Trudgill & Hannah, 1994; Tsukada, 2008).

According to the acoustic results, /eɪ/ indicates significant differences in F1 and F2 values and duration for each phase after training, possibly reflecting the fact that participants made an effort to differentiate between two-vowel combinations during their individual practice sessions. In addition, a comparison of the production of the diphthongs /eɪ/ and /e:/ after training reveals significant differences in relation to formant frequency and duration. The participants may have attempted to produce /eɪ/ and /e:/ differently. For /əʊ/, differences in some acoustic values between phases are not significant. The learners seem to have had difficulty with the production of this vowel. However, the duration can be considered the most prominent feature in their efforts. The vowel duration of /əʊ/ is shorter than that of /o:/ before training, but long after training. The learners may have emphasized longer duration to achieve diphthong quality.

The F1 and F2 values for /eə/ are not much different after training. Production of this diphthong shows significant duration before and after training as well as between phases. ROC indicates lesser formant movement after training. Vowel movement appears to be towards the back and center by the end of training. The results are likely to reflect the similarity of formant frequencies between /eə/ and /ɛ:/. Even though the duration of these two vowels is different, their F2 values are relatively close, with non-significant differences in values after training. For /eə/, the learners seem to pronounce this vowel with a monophthong quality more than in the case for /eɪ/ and /əʊ/. The learners may be producing /eə/ as the pure /ɛ:/ in Thai.

"Similar" features from L1 phones can help and interfere with L2 performance at the same time. While students participated in one month of individual practice on PRAAT, this may have been too short a period for pronunciation practice of some diphthongs. The Thai EFL learners may have required more time to differentiate /eɪ/, /əʊ/ and /eə/ from /e:/, /ɛ:/ and /o:/, and to eventually produce these three diphthongs using "New" phones. According to the acoustic results, the participants may achieve /eɪ/, followed by /eʊ/ and /eə/.

**The use of PRAAT in Thai EFL pronunciation practice**

PRAAT seems to have benefited the Thai learners in this study, helping them to improve their pronunciation, even though the program was primarily designed for speech analysis and synthesis. The acoustic results demonstrate some changes in the language production of the Thai EFL learners after their period of training on PRAAT. During the study, they were able to observe their performances directly via a visual spectrogram. This allowed them to notice the difference between diphthongs and monophthongs, helping to motivate them in their efforts to achieve accurate pronunciation of two-vowel combinations during their practice sessions. Training with a visual spectrogram can help EFL learners to improve their pronunciation, as has also been supported in previous studies involving EFL learners (Aramipoor & Gorjian, 2018; Gorjiana et al., 2013; Motohashi-Saigo & Hardison, 2009; Olson, 2014; Saito, 2007).

Most of the participants in this study tended to show positive feedback after their period of individual practice on PRAAT. They realized how useful the program was by helping them with their pronunciation training. PRAAT is likely to continue to have a positive impact on the Thai EFL learners in this study, reflecting arguments made in recent research by Osatananda and Thinchan (2021). Some participants may require more time for training on this program. Teachers need to provide methods that encourage learners to improve their pronunciation (Fraser, 2000).

All of the participants in this study were enrolled in a phonetics class during the period in which the research was conducted. They were able to gain some phonetic knowledge and learn how to operate PRAAT. However, some may have not been accustomed to individual practice or autonomous learning before they participated in this study. Therefore, they were all provided with a weekly in-class follow-up session and given two-hour supplementary outside of class during the week. It was important for the learners to be provided with clear instructions and assistance during the entire period of individual practice. This helped encourage them to participate effectively in their pronunciation practice sessions using the program provided.

## Conclusion and Recommendation

Duration is likely to have been the main parameter used by the Thai EFL learners in this study to assess improvements in their production of English diphthongs. Vowel length distinction in Thai seems to have affected the learners. The first vocalic element was likely to have been the most important phase in diphthong production. The Thai EFL learners struggled with the three English diphthongs /eɪ/, /əʊ/, and /eə/ due to the absence of these sounds in the Thai language. These diphthongs tended to be replaced by the monophthongs /eː/, /ɛː/, and /oː/ respectively. However, some features of diphthongs and monophthongs appeared to be different after the training. The learners may have required more time to practice and achieve accurate production of some particular diphthongs.

Based on the results of this study, it is possible to conclude that PRAAT can benefit Thai EFL learners and help them with their pronunciation practice through the use of the visual spectrogram. Learners may feel less stressed because they can manage their training schedule at their own pace. If they have received clear instructions and are provided with assistance, they will likely be able to participate in individual pronunciation practice sessions using any program that has been provided, PRAAT being one example.

Due to the COVID-19 pandemic lockdown, onsite recording after training could not be carried out. The recordings were made at home by the participants themselves in environments with different acoustic conditions. This factor explains the presence of strong background noise in certain audio files, even though participants had set the same sampling rate on PRAAT. These files were excluded from the analysis to help avoid misinterpretation.

Nowadays, Thai EFL learners are having to spend more time engaged in autonomous learning. The use of technology for pronunciation practice should be more widely promoted. In addition, further research could provide further insights into differences in the production of phonetic features by Thai EFL learners before and after practice with PRAAT. The training period could also be extended to longer than one month, and a perception test should also be conducted.

## Acknowledgment

I would like to express my sincere gratitude to the Faculty of Humanities, Srinakharinwirot University for a grant (560/2562) that allowed me to complete this research. Special thanks to the participants in this project.

## References

Aramipoor, S., & Gorjian, B. (2018). The effect of using PRAAT software on pre-intermediate EFL learners' supra segmental features. *Journal of Studies in Learning and Teaching English*, *7*(1), 13th ser., 47–68.

Arunsirot, N. (2020). Implementing the augmented reality technology to enhance English pronunciation of Thai EFL students. *KKU Research Journal (Graduate Studies) Humanities and Social Sciences*, *8*(3), 142–153. https://so04.tci-thaijo.org/index.php/gskkuhs/article/view/247821

Boersma, P., & Weenink, D. (2009). *Praat: Doing phonetics by computer [computer program]*. Version 5.1.05. http://www.praat.org./

Boersma, P. & Weenink, D. (2019). *Praat: doing phonetics by computer [computer program]*. Version 6.1.07. http://www.praat.org./

Brett, D. (2004). Computer generated feedback on vowel production by learners of English as a second language. *ReCALL, 16*(1), 103–113. https://doi.org/10.1017/s0958344004000813

Chapelle, C. (2004). Technology and second language learning: Expanding methods and agendas. *System, 32*(4), 593–601. https://doi.org/10.1016/j.system.2004.09.014.

Dee-in, W. (2006). The development of computer-assisted instruction lessons promoting oral skills for English problem sounds for students of English majoring in international communication. *Rajamangala University of Technology Lanna Phitsanulok Campus*. Thesis. M.ED. Uttaradit Rajabhat University.

Deerajviset, P. (2014). Technology in EFL teaching and learning in Thailand: An overview of research and issues. *Journal of Mekong Societies*, *10*(1), 71–112. https://so03.tci-thaijo.org/index.php/mekongjournal/article/view/17725

Deterding, D. (2000). Measurements of the /eɪ/ and /əʊ/ vowels of young English speakers in Singapore. In A. Brown, D. Deterding & E. L. Low (Eds.), *The English language in Singapore: Research on pronunciation*, (pp. 93–99). Singapore Association for Applied Linguistics.

Farooq, M., & Mahmood, D. A. (2017). Acoustic analysis of front vowels /ɛ/ and /æ/ in Pakistani Punjabi English. *International Journal of English Linguistics*, *8*(1), 219–229. https://doi.org/10.5539/ijel.v8n1p219

Flege, J. E. (1987). The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics, 15*(1), 47–65. https://doi.org/10.1016/S0095-4470(19)30537-6

Fouz-González, J. (2020). Using apps for pronunciation training: An empirical evaluation of the English file pronunciation app. *Language Learning & Technology*, *24*(1), 62–85.

https://www.researchgate.net/publication/333295662_Using_apps_for_pronunciation_training_An_empirical_evaluation_of_the_English_File_Pronunciation_app

Fraser, H. (2000). Coordinating improvement in pronunciation. In M. Celce-Murcia (Ed.) *Teaching English as a Second or Foreign Language*, (pp. 136–153). Newbury House.

Fry, D. B. (1979). *The physics of speech*. The Press Syndicate of the University of Cambridge.

Gorjiana, B., Hayati, A., & Pourkhonic, P. (2013). Using Praat software in teaching prosodic features to EFL learners. *Procedia - Social and Behavioral Sciences*, *84*, 34–40. https://www.sciencedirect.com/science/article/pii/S1877042813015723?via%3Dihub

Iadkert, K., & Hashim, A. (2020). An acoustical analysis of diphthongs in Thai English trajectory. *The 7th Symposium of International Language and Knowledge 2020*, 1–8. https://la.mahidol.ac.th/silk2020/wp-content/uploads/2020/12/Proceedings-SiLk-2020.pdf

Kelly, G. (2000). *How to teach pronunciation*. Longman.

Kriedler, C. W. (1989). *The pronunciation of English*. Basil Blackwell.

Ladefoged, P. (2006). *A course in phonetics* (5th ed.). Wadsworth Publishing.

Li, K. (2019). The application of Praat in English pronunciation teaching. *Proceedings of the 6th International Conference on Education, Language, Art and Inter-cultural Communication (ICELAIC 2019)*, 374–376. https://doi.org/10.2991/assehr.k.191217.150

Lightbown, P.M. & Spada, N. (2006). *How languages are learned*. Oxford University Press.

Motohashi-Saigo, M., & Hardison D.M. (2009). Acquisition of L2 Japanese geminates: Training with waveform displays. *Language Learning & Technology, 13*(2), 29–47.

Moxon, S. (2021). Exploring the effects of automated pronunciation evaluation on L2 students in Thailand. *IAFOR Journal of Education, 9*(3), 41-56. https://doi.org/10.22492/ije.9.3.03

Offerman H. M. & Olson D. J. (2016). Visual feedback and second language segmental production: *The generalizability of pronunciation gains. System*, *59*, 45–60. https://doi.org/10.1016/j.system.2016.03.003

Olson, D. J. (2014). Benefits of visual feedback on segmental production in the L2 classroom. *Language Learning & Technology*, *18*(3), 173–192. https://doi.org/10125/44389

Osatananda, V., & Thinchan, W. (2021). Using Praat for English pronunciation self-practice outside the classroom: Strengths, weaknesses, and its application. *LEARN Journal: Language Education and Acquisition Research Network, 14*(2). 372–396.

Pardede, P. (2018). Improving EFL students' English pronunciation by using the explicit teaching approach. *JET (Journal of English Teaching), 4*(3), 143. https://doi.org/10.33541/jet.v4i3.852

Pennock-Speck, B. (n.d.). RP (Received Pronunciation). *English accents*. https://www.uv.es/anglotic/accents_of_english/01/rp_received_pronunciation.html

Pillai, S. (2012). An instrumental analysis of English monophthongs produced by Thai EFL learners. *Partanika Journal of Social science and Humanities 20*(4). 1143–1160. https://www.academia.edu/2039446/Pillai_S_and_Salamae_W_2012_An_instrumental_analysis_of_English_monophthongs_produced_by_Thai_EFL_learners_Pertanika_Journal_of_Social_Sciences_and_Humanities_20_4_1143_1160

Priya M.L.S., Kumar P. (2020). Teaching phonetics to enhance pronunciation in an ESL classroom. *Journal of Critical Reviews, 7*(2), 669–672. https://doi.org/10.31838/jcr.07.02.121

Roach, P. (2004). British English: Received Pronunciation. *Journal of the International Phonetic Association, 34*(2), 239–245. DOI:10.1017/S0025100304001768

Roengpitya, R. (2002). Different durations of diphthongs in Thai: a new finding. *Proceedings of the Twenty-Eighth Annual Meeting of the Berkeley Linguistics Society: Special Session on Tibeto-Burman and Southeast Asian Linguistics.* 43–54. DOI:10.3765/bls.v28i1.1061

Saadah, F., & Ardi, H. (2020). The analysis of students' pronunciation error on English diphthong made by fifth semester of English language education program Universitas Negeri Padang. *Journal of English Language Teaching, 9*(1), 188–194. http://ejournal.unp.ac.id/index.php/jelt/article/view/107829

Sahatsathatsana, S. (2017). Pronunciation problems of Thai students learning English phonetics: A case study at Kalasin university. *Journal of Education, Mahasarakham University, 11*(4*)*. 67–84. http://edu.msu.ac.th/journal/home/journal_file/383.pdf

Saito, K. (2007). The influence of explicit phonetic instruction on pronunciation teaching in EFL settings: the case of English vowels and Japanese learners of English. *The Linguistics Journal 3*(3), 17–41. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.418.5662&rep=rep1&type=pdf#page=16

Sumbayak, D. M. (2009). The diphthongs: The obstacles for Indonesian speakers of English. *Lingua Cultura, 3*(2), 107-115. https://doi.org/10.21512/lc.v3i2.337

Tingsabadh, M., & Abramson, A. (1993). Thai. *Journal of the International Phonetic Association, 23*(1), 24–28. https://doi.org/10.1017/S0025100300004746

Trudgill, P., & Hannah, J.A. (1994). *International English: A guide to varieties of standard English* (3rd ed.). Edward Arnold.

Tsukada, K. (2008). An acoustic comparison of English monophthongs and diphthongs produced by Australian and Thai speakers. *English World-Wide, 29*, 194–211. https://doi.org/10.1075/eww.29.2.05tsu

Tsukada, K. (2009). Durational characteristics of English vowels produced by Japanese and Thai Second Language (L2) learners. *Australian Journal of Linguistics, 29*, 287-299. https://doi.org/10.1080/07268600902823144

Wilson, I. (2008). Using Praat and moodle for teaching segmental and suprasegmental pronunciation. *Proceedings of the 3rd International WorldCALL Conference: Using Technologies for Language Learning (WorldCALL 2008).* https://www.researchgate.net/publication/237734724_Using_Praat_and_Moodle_forTeaching_Segmental_and _Suprasegmental_Pronunciation

Wongsuriya, P. (2020). Improving the Thai students' ability in English pronunciation through mobile application. *Academic Journals, 15*(4). 175–184. https://academicjournals.org/journal/ERR/article-abstract/BE7011863359

Yangklang, W. (2013). Improving English stress and intonation pronunciation of the first-year students of Nakhon Ratchasima Rajabhat university through an e-learning. *Procedia - Social and Behavioral Sciences, 91*, 444–452. https://doi.org/10.1016/j.sbspro.2013.08.442