Selected Pronunciation Issues of South Vietnamese English

Ondrej Slowik, Doan Thi Hue Dung

Hong Bang University, Ho Chi Minh City, Vietnam

Email: ondra.slowik@seznam.cz

Abstract

In the process of English language acquisition by Vietnamese learners, speaking skills and pronunciation, in particular, tend to be the competencies most arduous to acquire. We assume that learners' weaknesses in speaking performance might be partially conditioned by the lack of confidence to express oneself verbally due to pronunciation difficulties. The objective of our article is to describe selected pronunciation issues of South Vietnamese speakers of English that we identified in our preliminary research in the form of a case study carried out on the students at the English Department of the Hong Bang International University in Ho Chi Minh City. We described the case study in great detail in the article by Slówik & Doan published in Vietnamese. Our task here was to utilize the findings of our case study and confirm whether it is possible to support our conclusions using data from computer speech analysis with the help of the software PRAAT. In order to do so, we had to compose a sub-corpus of 6 speakers of South Vietnamese English (3 males, 3 females, all in their early 20 s and born in Ho Chi Minh City). The results discussed in this article can be utilized for the purposes of applied linguistics or ELT education when working with (South) Vietnamese learners.

Keywords

Vietnamese English, Pronunciation, Foreign Accent, TEFL

1. Introduction

In the process of English language acquisition by Vietnamese learners, speaking skills in general and pronunciation, in particular, tend to be the competencies most arduous to acquire. This claim can be inferred from the classroom behavior of students and observing the interaction of Vietnamese speakers of English

with foreign educators, business associates, tourists, etc. as well as from the scores in various English proficiency tests where the assessment of speaking consistently reaches the lowest values. We assume that the generally poor performance in speaking examinations might be partially conditioned by the lack of confidence to express oneself verbally due to pronunciation difficulties. According to Lambert et al. (1960), "listeners naturally attribute social identity to speakers, and then judge those speakers in accordance with their stereotypes of the speaker's putative social group". Moreover, numerous studies show (e.g. (Giles, 1970)) that accent and pronunciation are key aspects of first-impression assessment i.e. listeners tend to ascribe character features and intellectual properties to the speaker based on his/her ability to enunciate distinctly in accordance with the pronunciation norms of the given language.

According to Cunningham (2009), "*The pronunciation of English presents severe challenges to Vietnamese-speaking learners. Not only is the sound system (…) very different from that of English, but there are also extremely limited opportunities for hearing and speaking English in Vietnam. In addition, there are limited resources available to teachers of English in Vietnam so teachers are likely to pass on their own English pronunciation to their students. University students of English are introduced to native-speaker models of English pronunciation but they do not often have the opportunity to speak with non-Vietnamese speakers of English.*"

The objective of our research is to describe selected pronunciation errors in English spoken by South Vietnamese speakers. We based our research on the article Slówik & Doan (2021) published in Vietnamese where we designed an experimental text containing lexemes prone to revealing pronunciation issues of Vietnamese English speakers. 100 freshmen from Hong Bang University English Department were asked to read the text out loud and their performance was then assessed by 2 examiners. The pronunciation issues revealed by this survey were then classified into three categories based on their level of linguistic stigmatization.

This article attempts to confirm whether these pronunciation issues can be detected using speech analysis software. In order to do so, we decided to record 8 speakers of South Vietnamese English reading excerpts from BBC News that we will proceed to analyze using the software PRAAT (Boersma & Weenink, 2020). If we gain reliable data supporting our results from the preliminary research, we will be able to visualize the pronunciation anomalies that the South Vietnamese learners utter and it will hopefully help us find a methodological remedy applicable in ELT.

2. Method

2.1. Preliminary Research

In the article Slówik & Doan (2021), we decided to use the British standard mainly because it is utilized in the HIU curriculum as well as for technical reasons be-

cause there is an extensive phonetic corpus focused on accent variation based on the recordings of BBC news that will enable us to conduct further research later on. According to Peter Roach (2006: p. 240), "the accent has been known for nearly a century as *Received Pronunciation*, or by its abbreviation, RP".

With the help of Hancock (2003), Cunningham (2009) and Singer (2012), we selected 17 potential pronunciation issues (listed in Table 1) that we expected might cause problems for Vietnamese speakers of English. About a third of the issues were chosen as a "control group" based on Hancock's observation that they constitute a problem for the majority of English learners even though judging from our own in-class experience and the works of Cunningham and Singer, we did not expect them to cause any problems to Vietnamese learners.

Every pronunciation issue was subsequently modelled into an experimental sentence. E.g. experimental sentence for issue No. 1.: *Peter put his big bag on the purple table behind the blue PC*; No. 16.: *The English language course costs six hundred dollars.* Finally, 100 first year students were asked to read the set of 17 experimental sentences while simultaneously being judged on the scale 1 - 5; 5 being the most fluent and correct while 1 the most erroneous. The assessment was carried out by the two authors of this article and the pronunciation issues were then categorized based on the extent they hinder comprehensibility of the speakers. Here we are going to address some of the most prominent ones that occurred most frequently and can cause the most problems with comprehension of the speakers.

2.1.1. Consonant Discrimination and Elision

Slówik & Doan (2021) identify two English consonant pairs causing problems to South Vietnamese speakers: /s/ x /ʃ/ and /p/ x /b/. Standard Vietnamese contains neither /p/ nor /ʃ/ as a natural phoneme. /p/ exists in loanwords such as *pin* or

Table 1. Potential pronunciation issues for Vietnamese speakers of English. They are mostly concerned with discrimination of two sounds or ability to produce the desired sound correctly. 8 and 9 deal with aspiration of stressed voiceless plosives /p, t, k/ and its lack after /s/. Elision stands for simplification or reduction of consonants clusters. Tone captures the interference of Vietnamese lexical tones over English syllables (Slówik & Doan, 2021).

1.	p x b	10.	tr
2.	s x ʃ	11.	ʊ x u:
3.	s x z	12.	aɪ/eɪ/ɔɪ
4.	tʃ	13.	æ
5.	dʒ	14.	ɪx i:
6.	th	15.	ʌ
7.	v x w	16.	*elision*
8.	*aspiration*	17.	*tone*
9.	*s (ptk)*		

pê đê and the grapheme *s* can be pronounced as /ʃ/ in certain dialects (Thompson, 1965). However, despite the existence of these sounds in the Vietnamese sound inventory, many of the speakers struggle with their correct pronunciation. They tend to pronounce /p/ as /b/ i.e. not only omitting aspiration but also voicing it. This can be remedied by reminding the students that they should be able to pronounce /p/ even in their native language and accentuating the aspiration element. In the case of /s/ x /ʃ/ discrimination, a large number of the speakers were unable to pronounce /ʃ/ replacing it with /s/ hence rendering minimal pairs e.g. *sip* x *ship* to sound like /sɪp/.

The *th* sound both voiced /ð/ and voiceless /θ/ often end up being realized mostly as a heavily aspirated voiceless dental stop [tʰ], a sound native to the Vietnamese language. Although this is a substandard variation, it does not affect comprehensibility very much as /ð/ and /θ/ are not mutually distinctive and the heavy aspiration distinguishes the sound from /t/ and /d/. Moreover, most Irish dialects of English actually replace /ð/ and /θ/ with dental stops (Hickey, 2004) and foreigners tend to struggle with the dental fricatives in general and so the native listeners seem to be more forgiving when they hear the sound mispronounced.

There are no /dʒ/ and /tr/ sounds in standard Vietnamese (although Thompson claimed that the dialects around the city of Nha Trang can at times pronounce the /tɕ/ sound as /tr/), and a substantial number of students realized /tr/ as [tɕ], which leads to the lexemes *train-chain* both sound as [tɕeɪn] or rather [tɕe:n]. /dʒ/ was often simplified to [ʒ] or [z]. However, there are very few minimal pairs of /dʒ/ and /ʒ/ or /z/ in English and those that exist such as *juice* /dʒus/ and *Zeus* /zus/ are hardly ever confused.

It is quite characteristic for the English language to place more consonants together and create consonant clusters. The clusters can be placed syllable-initially as well as finally. Roach (2006) distinguishes two-consonant and three-consonant initial consonant clusters and two-, three- and four-consonant final clusters. There are no consonant clusters in Vietnamese and thus a consonant can only occur syllable-initially or syllable-finally. Furthermore, it is always followed or preceded by a vowel. Vietnamese phonotactics hence does not allow for consonant clusters and therefore the Vietnamese tend to simplify them when speaking English as well. Moreover, Vietnamese syllables are semantically prominent at the beginning. In terms of consonants, Vietnamese syllables can only end in nasals or unreleased voiceless plosives. This phenomenon constitutes a major source of confusion as words like *next* can often be mispronounced as *neck, nest, net*. As mentioned above, except for nasals, there are no other means to close syllables other than by using voiceless plosives. Vietnamese speakers of English therefore often leave the syllables open (*like* pronounced as [lai]) or replace syllable-final consonants with voiceless plosives (*English* as [ɪŋglɪt] or knife as [naɪp]).

2.1.2. Vowel Production and Discrimination

Textbooks usually distinguish the two English vocalic pairs /ʊ/ x /u:/ and /ɪ/ x /i:/ based on both quantity as well as quality i.e. their length and position on the

vowel diagram (see **Figure 1**). However, "it should be understood that the terms "long" and "short" should be seen in relative terms: the vowels of both classes are subject to the lengthening and shortening effects found in English, with the result that a "short" vowel may, in some contexts, be longer than a "long" vowel in a different context. The length mark /:/ is used to mark the long vowels, though this is actually redundant since the vowel symbols already successfully distinguish each vowel from every other (Roach, 2006: p. 241)."

Vietnamese speakers of English manifest great difficulties when differentiating between /ɪ/ x /iː/ and ʊ x uː. **Figure 1** shows us that Vietnamese /u/ and /i/ are situated in positions similar to the English /uː/ and /iː/. The English vowels /ʊ/ and /ɪ/ that are lower and laxer do not have any Vietnamese counterparts, which is likely the reason for this pronunciation issue. Moreover, due to the combination of pre-fortis clipping, a physiological phenomenon decreasing vowel duration before voiceless consonants (Wells, 2005), and the tendency of Vietnamese learners to replace all syllable-final consonants with voiceless plosives, the distinction within the pairs becomes even more complicated.

The open low vowel /ʌ/ is also not present in the Vietnamese vocalic inventory. Similar to other non-native speakers (Volín, 2005), Vietnamese speakers tend to lump the lacking sounds with those that they know from their mother tongue. As /ʌ/ is located between the Vietnamese /ə/ and /a/, the Vietnamese speakers lump it together with /ə/ and so the English word *but* sounds similar to the Vietnamese *bất* [bət35]. This is also caused by the fact that /ʌ/ occurs only in closed syllables that Vietnamese learners often pronounce with the rising tone sắc (see 2.1.3).

The research also suggested that South Vietnamese speakers tend to fail in the production of the final /ʊ/ and /ɪ/ in English diphthongs and therefore the English word *time* in their rendering resembles the Vietnamese word *tham* [thaːm55] or *came* sounds more like [keːm55]. Vietnameseutilizes only falling centering diphthongs, whereas English speakers can draw from an inventory of 6 rising diphthongs (**Figure 2**). Therefore, Vietnamese speakers of English struggle with the pronunciation of English rising diphthongs.

2.1.3. Tonal Interference

Apart from elision described above, tonal interference turned out to be the most
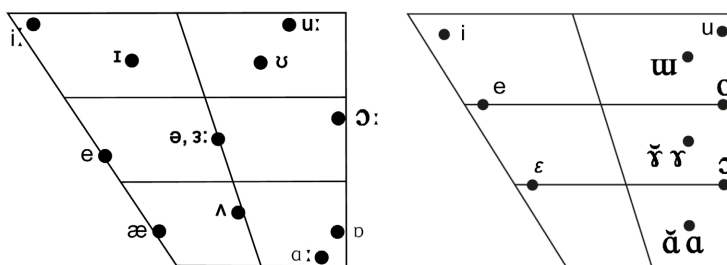

Figure 1. Vowel diagram of British English (left (Roach, 2006)) and standard Vietnamese (right (Kirby, 2011)). Vietnamese /u/ and /i/ are positioned very similarly to English /uː/ and /iː/. English /ʊ/ and /ɪ/ lack Vietnamese counterparts. The English vowel /ʌ/ is also the case as there is no vowel in the space between /ə/ and /a/.
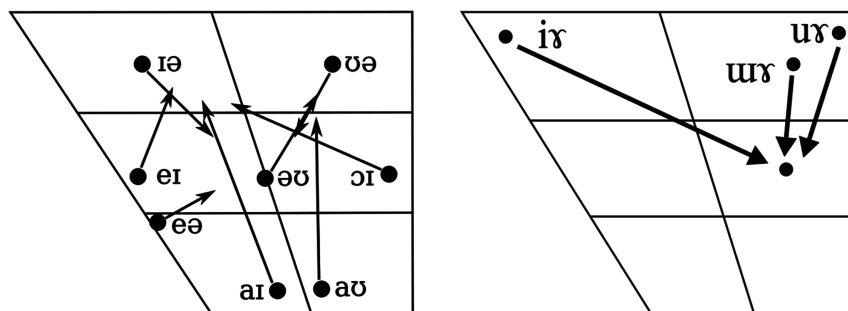
Figure 2. Diphthongs in British English (left (Roach, 2006)) and standard Vietnamese (right (Kirby, 2011)). Vietnamese utilizes only falling centering diphthongs, whereas English speakers can draw from an inventory of 6 rising diphthongs.

prominent pronunciation issue revealed in the results of the preliminary research from Slówik & Doan (2021). Due to rules of phonotactic and graphotactic, Vietnamese speakers are prone to reading the English open syllables with high level tone *ngang* and closed syllables with the rising *sắc*. Explanation of this phenomenon might lie in the fact that majority of borrowings into Vietnamese are pronounced with *ngang* (Pham, 2003) and the only type of syllable-final consonants other than nasals are unreleased plosives that restrict their tone affiliation to *sắc* and *nặng*. It can therefore be challenging for the Vietnamese speakers to rid themselves of the reflex. According to Cruttenden (1997), the basic natural intonation pattern in English is a mild decrease with a heavier fall at the end in terms of statements or a rise in certain types of questions. High flat intonation on open syllables and sharp rises on syllables with plosives or fricatives in their coda could be a reason for the "choppiness" of Vietnamese English described by Cunningham (2009).

2.2. Material Recording, Processing and Analysis

Our target was to create a small sub-corpus of spoken South Vietnamese English. We managed to record 8 speakers for approximately 9 minutes each. The speakers were labelled by five capital letters. The first letter A-D stands for the order of speakers, after that SVE stands for Vietnamese English and the last letter F/M represents gender. All the speakers were selected among the students of the 1st year at the HIU English Department, 4 males and 4 females. The main criteria were the origin from the Ho Chi Minh City and good English reading fluency but with prominent Vietnamese accent.

The speakers were asked to read BBC news (each speaker about 1000 words) from the database of newsreaders used by many phonetic departments across the globe. We used texts labelled AMA and JLA (AMA-Alice Moss, JLA-Jackie Leonard). Each text comprised approximately 500 words. All recordings were captured using a recording device with 4 microphones allowing stereo recording, sampling frequency up to 48,000 Hz, MP3 (up to 640 kb)/WAV (16/32bit) format and a button to adjust sound input. The format WAV 16 bit and sampling frequency 32,000 Hz was set as default setting for all the recordings.

The recordings were transferred from the device on a computer and labelled according to the abbreviations of BBC presenters and gender of the speaker. Adobe Audition software was used to convert the recordings to mono and then they were subjected to the analysis in the phonetic software PRAAT and Forced Aligner. As the last step of data preparation, the work of Forced Aligner had to be manually checked in the PRAAT files.

Finally, we used PRAAT scripts as well as manual selection methods to localize specific instances of the pronunciation issues of South Vietnamese speakers of English that were previously identified based on the preliminary research and described in detail in Section 2.1. In order to do so, all sounds were annotated to Text Grids and segmented into lexemes and phonemes in two separate layers.

3. Results

3.1. Consonant Issues

Among the consonant issues, we managed to identify the substandard pronunciation of /θ/ (see **Figure 3**), insufficient discrimination of /s/ x /ʃ/ (**Figure 4**) and consonant cluster simplification or their complete elision (**Figure 5** and **Figure 6**).

In **Figure 3**, we can clearly see from both the spectrogram as well as the oscillogram that the speaker ASVEF pronounced the fricative /θ/ as an affricate [tʃ]. Alternatively, Vietnamese speakers can also pronounce it as [tʰ] because [tʰ] is a natural item in the inventory of Vietnamese consonants. As standard British English does not utilize the [tʰ] sound, this issue so typical for Vietnamese speakers does not constitute a large problem in intelligibility and it only requires a short period of getting used to by the listener.
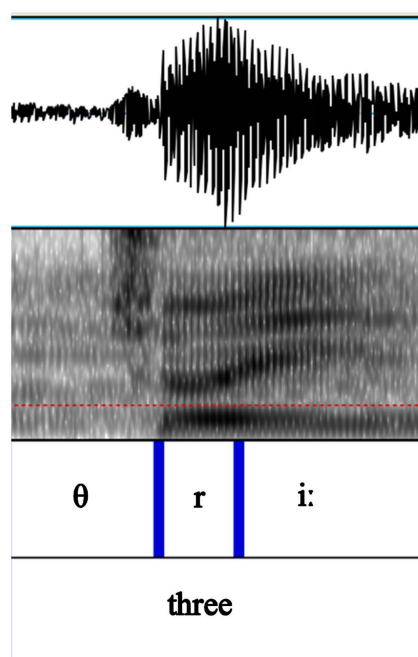

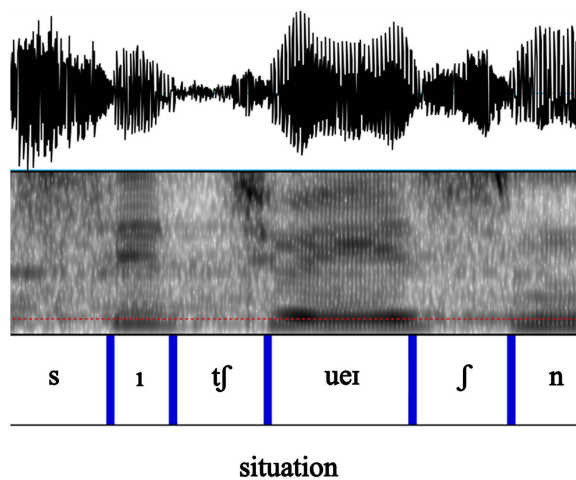**Figure 3.** Speaker ASVEF realized the fricative /θ/ as an affricate [tʃ].

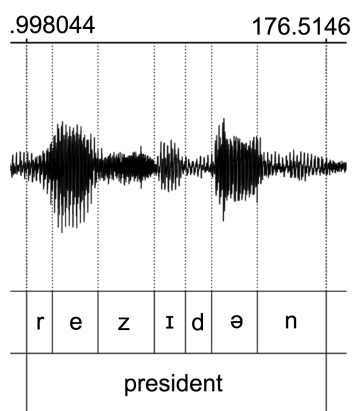**Figure 4.** The speaker BSVEF pronounced the word "situation" with an /s/ instead of /ʃ/ at the end.

Figure 5. Speaker CSVEM pronouncing the word "president" without the initial /p/ and final /t/.
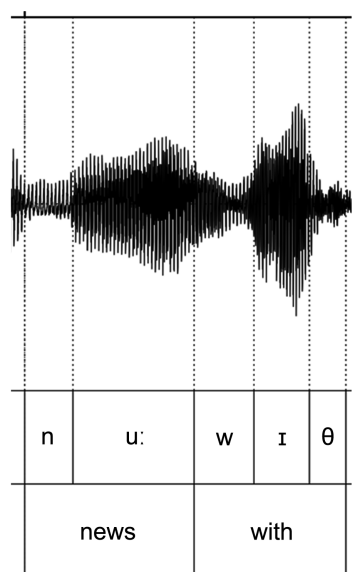

Figure 6. Speaker BSVEM demonstrates the omission of a final /s/ consonant.

In **Figure 4**, the speaker BSVEF pronounced the word "situation" with an /s/ instead of /ʃ/ in the final syllabe. The sound /ʃ/ does not occur in standard Vietnamese and it is becoming a rarity in the Southern dialect as well. The sound /tʃ/, on the other hand, belongs to the consonant itinerary of the standard as well as Southern Vietnamese and therefore it is pronounced clearly in English as well. The inability to distinguish /s/ and /ʃ/ has been a frequent issue throughout our recorded materials.

Figure 5 illustrates elision or consonant cluster simplification when the speaker CSVEM pronounces the word "president" without the initial /p/ and final /t/. Speaker BSVEM in **Figure 6** demonstrates the omission of a final /s/ consonant. This phenomenon leads to many grammar-related misunderstandings due to the fact that it affects the 3rd person singular nouns, possessives, plurals and contractions. Therefore, it is often difficult to determine the nature of the error i.e. whether it is a matter of pronunciation or grammar. Consonant omission (especially /s/) and consonant cluster simplification are two features of (South) Vietnamese English that were prominent throughout our recorded material as well as during the reading sessions in stage 1.

3.2. Vocalic Issues

Regarding vowels, most speakers displayed insufficient discrimination of /ɪ/ x /iː/, which is often connected to the issue of pronouncing closed syllables with rising intonation (see 3.3.), and a substandard pronunciation of the vowel /ʌ/ (**Figure 7**) with a significantly more open vowel as well as a clear rise in the vowel pitch. Vietnamese phonotactics only allows for two tones in syllables closed with plosives (rising and low), which leads to the interference into Vietnamese English and most of the closed syllables with final plosives or fricatives experience a significantly rising pitch often leading to the wrong assumption that the speaker is in fact posing a question.
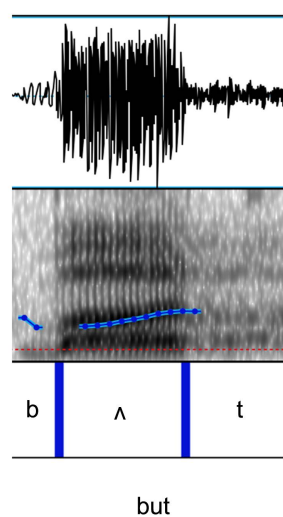

Figure 7. Speaker ASVEF pronouncing the word "but" with a rising pitch and a significantly more open vowel.

In **Figure 8**, Speaker ASVEF pronounced the diphthong /eɪ/ without the final element hence it sounded like a single vowel /e/ with a longer duration. Vietnamese utilizes 3 diphthongs but they are all falling/centring (see **Figure 2**) and therefore the articulation of a rising diphthong constitutes a hindrance for Vietnamese speakers of English. Not being able to pronounce the final element of rising diphthongs can also be classified as one of the key features of (South) Vietnamese English pronunciation.

3.3. Tonal Interference

Figure 9 illustrates how pronunciation rules from Vietnamese manifest themselves in the Vietnamese English. As described in 2.1.3., Vietnamese open *ngang* and syllables closed with plosives or fricatives with rising *sắc*. This leads to the high pitch of the open syllables and rising pitch of the closed syllables in
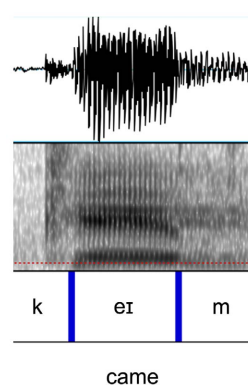

Figure 8. The diphthong /eɪ/ pronounced without the final element hence it sounded like a single vowel /e/ with a longer duration.
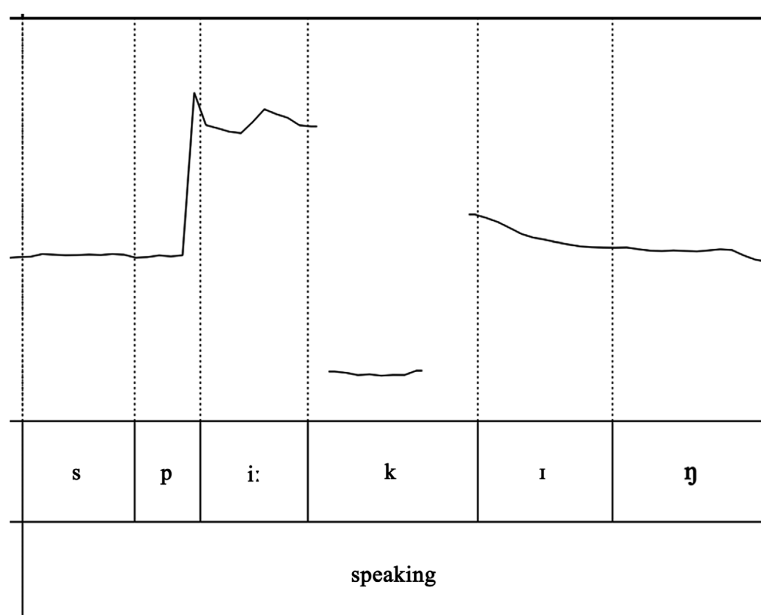

Figure 9. Rising pitch in the closed syllable "speak-" illustrates the fact that it is rather difficult for the (South) Vietnamese speaker not to raise voice in closed syllables.

syllables in borrowed words tend to be pronounced using the high flat tone (South) Vietnamese English. The word "speaking" is pronounced with a significant rise in pitch on the syllable "speak" that is very unnatural for native English. This phenomenon also leads to problems in proper distinction between /ɪ/ and /iː/ as the long vowel also tends to be higher in pitch (Roach, 2006).

4. Conclusion

Based on the findings of our preliminary perception experiment that identified three major types of pronunciation difficulties encountered by South Vietnamese speakers of English: 1) Consonant discrimination and elision; 2) Vowel production and discrimination; 3) Tonal interference, we were able to confirm most of these difficulties using the PRAAT speech analysis software. Namely, we found numerous instances of /s/ x /ʃ/ discrimination problems, articulation of /ð/ and /θ/, simplification of consonant clusters or complete consonant elision. In terms of vowel production, we presented examples of insufficient /ɪ/ x /iː/ discrimination, substandard /ʌ/ articulation and omission of the final /ɪ/ segment of English diphthongs. The PRAAT function of pitch contour mapping allowed us to notice that the recorded speakers quite frequently use high flat or sharply rising intonation in a manner very uncommon for natural English intonation.

The results of this study provide us with a rather clear portrait of the South Vietnamese English pronunciation issues and we are hoping that our findings might enable applied linguists and English teachers to gain deeper awareness with respect to their Vietnamese students' accents and subsequently they will be able to come up with effective methods to alleviate the most stigmatizing features of their pronunciation. Moreover, we also achieved to gather a substantial amount of spoken language recordings that can be effectively utilized in our future research endeavours.

Our research opened a number of opportunities for both future research as well as for implementing the findings into applied methodology in teaching pronunciation. Especially the issues of consonant elision and pitch rising in closed syllables are two topics that could be analysed in greater detail in the future. In terms of pronunciation teaching methodology, we would like to devise written guidelines for teachers of phonics operating in Vietnam.

Funded

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

Boersma, P., & Weenink, D. (2020). Praat: Doing Phonetics by Computer. https://www.fon.hum.uva.nl/praat/

Cruttenden, A. (1997). *Intonation*. Cambridge University Press.
https://doi.org/10.1017/CBO9781139166973

Cunningham, U. (2009). Phonetic Correlates of Unintelligibility in Vietnamese-Accented English. In *Proceedings*, *Fonetik 2009*. University of Stockholm.

Giles, H. (1970). Evaluative Reactions to Accents. *Educational Review, 22,* 211-227.
https://doi.org/10.1080/0013191700220301

Hancock, M. (2003). *English Pronunciation in Use*. Cambridge University Press.

Hickey, R. (2004). The Phonology of Irish English. In B. Kortmann, et al. (eds.), *Handbook of Varieties of English. Volume 1: Phonology* (pp. 68-97). Mouton de Gruyter.

Kirby, J. P. (2011). Vietnamese (Hanoi Vietnamese). *Journal of the International Phonetic Association, 41,* 381-392. https://doi.org/10.1017/S0025100311000181

Lambert, W. E., Hodgson, R. C., Gardner, R. C., & Fillenbaum, S. (1960). Evaluational Reactions to Spoken Languages. *Journal of Abnormal and Social Psychology,* 60, 44-51.
https://doi.org/10.1037/h0044430

Pham, A. H. (2003). *Vietnamese Tone: A New Analysis*. Routlege.
https://doi.org/10.4324/9780203500088

Roach, P. (2006). *English Phonetics and Phonology* (3rd ed.). Cambridge University Press.

Singer, E. (2012). *Vietnamese English*.
http://www.yorku.ca/earmstro/asia/Vietnamese_accent.pdf

Slówik, O., & Doan, H. D. (2021). Pronunciation Issues of Learner English: A Hong Bang International University Case Study. *HIU Scientific Journal*, 1-10.

Slówik, O., & Volín J. (2020). *Tone in Vietnamese Metropoles*. Karolinum.

Thompson, L. C. (1965). *A Vietnamese Grammar*. University of Washington Press.

Volín, J. (2005). *IPA-Based Transcription for Czech Students of English*. Karolinum.

Wells, J. C. (2005). *Longman Pronunciation Dictionary*. Pearson Education Ltd.