

PERCEPTION OF ENGLISH LISTENERS ON VIETNAMESE VOWELS CONTRASTING IN ROUNDING

Tran Thi Hai Yen*

Thai Nguyen University

Received: 28/07/2020; Revised: 20/09/2020; Accepted: 26/04/2021

Abstract: This study investigated into how English listeners from Midland U.S. perceived Vietnamese back vowels contrasting in rounding. The words were produced by Northern and Central Vietnamese speakers. The results showed that Vietnamese vowels [o] and [u] was assimilated to English [oo] (77%) and [u] (69%) respectively. [ɤ] and [ɯ] were not assimilated to any English vowels. This suggested that Midwest U.S. speakers tended to use vowel height to map between native vs. non-native vowels while Southern U.S. speakers in the only previous study (Shport, 2019) tended to use vowel rounding. Dialect effects were found for Northern Vietnamese [ɯ] was assimilated to [u] (61%) while Central Vietnamese [ɯ] was perceived the most as [ɔ] (39%). Thus, the perception of non-native sounds can be affected by the dialects of both speakers and listeners. The study suggests that English teachers should be aware of this issue to accommodate appropriate teaching strategies in classroom.

Key words: Perception assimilation model, second language acquisition, second language perception, Vietnamese vowels

1. Introduction

Although naïve listeners' perception on non-native sounds has been well-studied (Levy, 2005; Simon, Debaene&Herreweghe, 2015; Shport, 2019), how listeners perceive back vowels contrasting in rounding is limited. Shport (2019) was the only first to study how English speakers perceived Vietnamese back vowels contrasting in rounding. Shport (2019) recruited English speakers with Southern U.S. dialect (SUSE) and a Vietnamese speaker with Central Vietnamese. But both Vietnamese and U.S. English have many regional dialects with different vowel characteristics. Thus, more research into how listeners' and speakers' dialects may affect this perception is needed. This study is a replication of Shport (2019) with two contributions regarding this issue. First, the speakers in this study were from both Central and Northern Vietnam. Kirby (2011) claimed that Northern Vietnamese [ɯ] was more mid-centralized than Central Vietnamese [ɯ]. Hence, how Northern Vietnamese [ɯ] was perceived compared with Central Vietnamese [ɯ] was a potential issue needing more in-depth research. Second, the listeners in this study were from the Midland U.S. (Kansas). English speakers from the Southern U.S. exhibited Southern Vowel Shift: the fronting of the back vowels /u/ and /oo/. Speakers from Midland U.S., on the other hand, did not necessarily exhibit the same process (Clopper, Pisoni& Jong, 2005). Regarding this, how these two vowels might yield different assimilation patterns is further explored in the current study.

* Email: tranhaiyen.sfl@tnu.edu.vn

2. Literature review

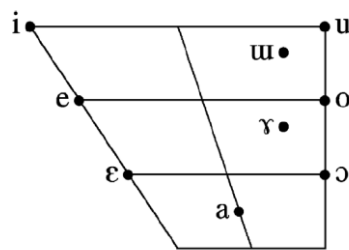
2.1. Perceptual Assimilation Model (PAM):

PAM (Best, 1995) predicts that naïve listeners perceive non-native contrasts based on articulatory/acoustic similarity with the native phones. If the non-native phone is perceived as ‘similar’ to L1 phones, the non-native phone will be ‘assimilated’ to the first language (L1) phone. As a result, listeners are unable to perceive the similar non-native phone correctly. PAM predicts five assimilation patterns:

- (1) Two-Categories (TC): each non-native phone is assimilated to a different L1 phone. For example, Spanish listeners should be able to discriminate English /ɪ/ and /ʊ/ accurately because these vowels are assimilated to Spanish /e/ and /o/ respectively. This assimilation pattern predicts excellent discrimination.
- (2) Category-Goodness (CG): two non-native phones are assimilated to the same L1 phone but how similar (goodness-of-fit ratings) they are to the L1 phone are different. For example, Spanish listeners should show moderate discrimination of English /ɪ/ and /eɪ/ because although both are assimilated to Spanish /e/, English /eɪ/ is more similar to Spanish /e/ than English /ɪ/. This assimilation pattern predicts moderate discrimination.
- (3) Single-Category (SC): two non-native phones are assimilated to the same L1 phone and they are equally different (goodness-of-fit ratings) from the L1 phone. For example, Spanish listeners should have difficulty discriminating English /ʌ/ and /ɑ/ because both are assimilated to Spanish /a/ and both are equally bad exemplars of Spanish /a/. This assimilation pattern predicts poor discrimination.
- (4) Uncategorized-Uncategorized (U-U): both non-native phones fall in the phonetic space but none of them are assimilated to any L1 phone. For example, Thai vowels [ɯ]-[ɤ] are categorized as English [ʌ] by only less than 70% of listeners so they are not assimilated to any English vowel. This assimilation pattern predicts from poor to good discrimination.
- (5) Uncategorized-Categorized (U-C): one non-native phone is not assimilated to any L1 phone while the other is. For example, Spanish listeners should not categorize English /u/ as any Spanish vowel, and should be able to discriminate between English /u/ and any other English vowel. This assimilation pattern predicts good discrimination.

Given the assimilation patterns above, the following order of accuracy discrimination is posited from the easiest to the most difficult: TC = UC > CG > UU > SC. To test the predictions of PAM, Levy (2005), Simon, Debaene & Herreweghe (2015) and Shport (2019) use two experiments: the first categorization experiment serves as predicted assimilation patterns and the second discrimination experiment tests naïve listeners’ discrimination of non-native phones.

Shport (2019) was the first study to examine how Southern U.S. English speakers (SUSE dialect) perceived the rounding contrast in Vietnamese back vowel pairs [u]-[o], [ɯ]-[u], [ɯ]-[ɤ], and [o]-[ɤ] according to PAM framework (Figure 1). The Vietnamese vowel pairs [o]-[ɤ], [u]-[ɯ] contrast in rounding; [u]-[o], [ɯ]-[ɤ] contrast in height.



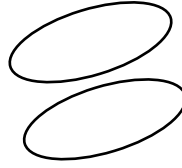


Figure 1. Vietnamese monothong centroids in a schematic F1-F2 space (Kirby, 2011)

In the first vowel categorization experiment, 49 English speakers listened to the four Vietnamese vowels and chose one from seven English vowels /u oʊ ʊ ɔ ʌ ɜ ɑ/ that was the most similar to the Vietnamese vowels they just heard. By calculating the percentage of response from participants, assimilation patterns could be established. When the 50% threshold was used as categorization criteria, the results of the first experiment basically suggested that Vietnamese [ɤ], and [u o] were perceived the most as English [ʌ] (70%) and [oʊ] (58% for [u] and 59% for [o]) respectively. [u] did not yield any dominant categorization (22% for [u], 32% for [ʊ], 26% for [ʌ]). The results from the first experiment predicted the following discrimination ability from the easiest to the most difficult: [o-ɤ] = [u-ɤ] = [u-u] > [u-o] as the assimilation patterns were UC = TC > SC accordingly. These predictions were tested in the second vowel discrimination experiment. In the discrimination experiment, English speakers made odd-man-out judgements for words with the Vietnamese vowels [u], [o], [u], [ɤ] presented in triads. In the experiment, participants heard three Vietnamese words including two words with the same vowel and one with a different vowel. The words with the same vowels were not the same productions: e.g. [tu1]-[tu1]-[tu2]. Then, participants selected the words with the different vowel compared with the other two. The second experiment basically supported the predictions with the observed ranking of accuracy: [o-ɤ] (90%) = [u-ɤ] (89%) > [u-u] (83%) = [u-o] (84%). The authors claimed that the unpredicted accuracy in [u-u] contrast actually could be predicted when the above-chance level (14.3%) categorization criterion was used. Overall, PAM could predict the speech perception in naïve listeners quite accurately.

1.2. Research questions and hypotheses

This experiment aimed to find out how English speakers from the Midland (Midwest) U.S. categorized Vietnamese vowels [u o u ɤ] to English vowels [u oʊ ʊ ɔ ʌ ɜ ɑ]. The following research questions and hypotheses were tested:

- How did English speakers from the Midwest U.S. categorize Vietnamese vowels into English vowels regarding both dialects? With the Southern Vowel Shift mentioned above, there might be a difference between this study and Shport (2019).
- How did English speakers from the Midwest U.S. categorize Vietnamese vowels into English vowels for Central Vietnamese dialect? If the Midwest U.S. speakers used the same acoustic cues (with Southern U.S. speakers) to map between native and non-native vowels, the assimilation patterns should be similar to Shport (2019).
- How did English speakers from the Midwest U.S. categorize Vietnamese vowels into English vowels for Northern Vietnamese dialect? Regarding the claim from Kirby (2011) that [u] was more mid-centralized for Northern Vietnamese dialect, this vowel may be perceived differently from Central Vietnamese dialect.

3. Methodology

3.1. Participants

For the production of Vietnamese stimuli, there were two female speakers who were native speakers of Vietnamese. One speaker was 26 years old and the other was 27 years old. These speakers reported Central and Northern Vietnamese as their dominant dialect.

For the vowel categorization experiment, participants were 21 (5 males, 16 females, mean age of 19.5 years old) native English speakers. They were all naïve to Vietnamese and were students at the University of Kansas, U.S. Their dialect was defined as Midwest (Kansas) American English. The difference between the number of speakers and listeners (2 vs. 21) may not cause any potential issue as discussed in Shport (2019).

3.2. Stimuli

The stimuli in this experiment were similar with the experiment of Shport (2019). For each of the Vietnamese vowel [u o u ɤ], there were two words (Table 1 and 2). Among the two words, one of them had level tone (e.g. *tu*) and the other had falling tone (e.g. *tù*). The words had CV structure in which the first consonants were always the voiceless dental stop [t] and the vowel was one of the above. Each word was produced five times by the speakers. The total number of the words for each dialect was forty: four vowels x two tones x five repetitions. The total number of tokens for both dialects was eighty. For the Vietnamese vowel production, the Northern Vietnamese speaker was recorded in anechoic chamber using a Marantz PMD 671 solid state recorder and an ElectroVoice N/D 767a microphone (16-bit resolution). The Central Vietnamese speaker was recorded in a quiet room using a mobile phone. The speakers produced the words with normal speaking rate.

3.3. Data analysis

To examine the Vietnamese vowel characteristics, duration, the fundamental (F0), first (F1), second (F2) and third formant frequencies (F3) of the Vietnamese vowels were measured (Table 1 and 2) and averaged across five repetitions. The frequencies were measured in the midpoints of the vowels. Although both English and Vietnamese did not distinguish long vs. short vowels, vowel duration was also measured because this could bias listeners in perceiving long vowels as tense and short vowels as lax. All the measurements were done in Praat software (<http://www.fon.hum.uva.nl/praat/>). Particularly, the wave form of each vowel was illustrated on Praat. Then, the researcher could measure its duration, F0, F1, F2 and F3 in the midpoint of the vowel. An example can be found in Figure 2.

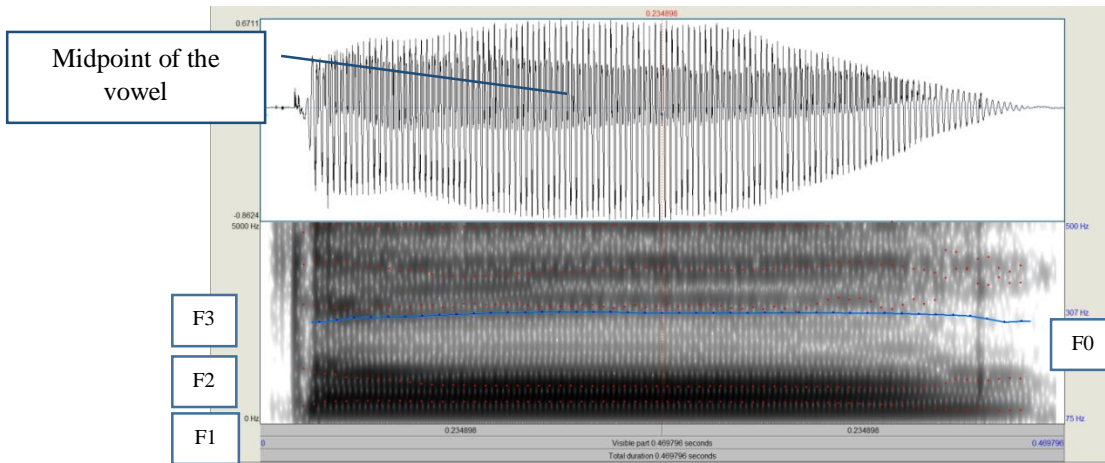


Figure 2. The waveform of the word token 'tô' with marked F1, F2, F3 and F0

To compare the characteristics of English vowels vs. Vietnamese vowel, a vowel space is plotted in section 4.2. (Figure 3). The data of English vowels were from Hillenbrand, Getty, Clark, and Wheeler (1995).

To establish the categorization patterns, the response proportion of each English vowel was calculated. This experiment used 50% threshold to determine assimilation patterns as in Shport (2019).

3.4. Data collection procedures

In the experiment, participants first completed a practice block with 8 Vietnamese words. The words of the practice block were the same with the test items: CV structure with a voiceless dental stop [t] preceding one of the vowels [u o u ɤ]. For each of the CV pair, there was one word with the level tone and one with the falling tone. The practice items were different productions from the test items and were produced by the speaker from Northern Vietnam. After the practice trials, participants completed the test block (80 items). In the experiment, participants first listened to a Vietnamese word, then they saw a list of seven English words on the screen. The words were GOOSE, GOAT, HAWK, PUT, BUS, POT and NURSE (the same with Shport, 2019). These words corresponded with the seven English vowel /u/, /oo/, /ɔ/, /ʊ/, /ʌ/, /ɑ/ and /ɜ/ respectively. Participants were instructed to click the mouse to the one English word that had the vowel the most similar with the vowel of the Vietnamese word they just heard. Finally, participants rated the token for goodness-of-fit of the English word to see how good the Vietnamese vowel as an example of the English vowel. The rate was from 1 (very bad), 2 (moderately bad), 3 (slightly bad), 4 (average), 5 (slightly good), 6 (moderately good), to 7 (very good). After that, participants pressed 'ok' button to move on to the next trial. The inter-stimulus interval was one second. Participants had unlimited time to provide their responses. All the stimuli were presented aurally via headphone on Paradigm. The stimuli were randomized across participants to avoid order effect. Participants were also instructed to make the best guess if they could not make a decision.

4. Findings and discussion

4.1. Vowel characteristics of Central and Northern Vietnamese vowels

The specific characteristics of the vowels produced by the two Vietnamese speakers are summarized in Table 1 and 2 below. The characteristics included mean vowel duration, F0, F1, F2 and F3 measured in the midpoints of the vowels. Some important generalizations of the characteristics are as follow. First, the duration of the words with falling tone (451ms) was always longer than the words with level tone (370ms). This effect was particularly strong in the Central speaker (456ms for words with falling tone and 334ms for words with level tone) compared with the Northern speaker (445ms for words with falling tone and 406ms for words with level tone). The mean durations of the words were shaded in the tables. This suggested that participants may perceive the vowels with falling tone as English tense vowels and the vowels with level tone as English lax vowels. Tone effect was not discussed in this paper because of its limited scope. Second, mean F1, F2, F3 in the vowels produced by the Central speaker most of the time were higher than these in the vowels produced by the Northern speaker. The plotting of F1-F2 in Figure 2 suggested that the difference in F1 and F2 in the production of two dialects was not very remarkable with the exception of [u]. Central Vietnamese [u] had the biggest difference in F1(82Hz) and F2(303Hz) compared with Northern Vietnamese [u].

Table 1. Vowel characteristics in eight Vietnamese words produced by Central Vietnamese speaker

Vowel	Word	Tone	F0 (Hz)	Duration (ms)	F1 (Hz)	F2 (Hz)	F3 (Hz)
[u]	<i>Tu</i> ‘knock up’	Level	293	289	367	909	3601
	<i>Tù</i> ‘prison’	Falling	218	444	459	975	3399
	Mean		255	366	413	942	3500
[ɯ]	<i>Tư</i> ‘four’	Level	296	322	406	1740	3682
	<i>từ</i> ‘word’	Falling	220	470	491	1705	3258
	Mean		258	396	449	1723	3470
[o]	<i>Tô</i> ‘bowl’	Level	275	331	633	929	3545
	<i>tô</i> non word	Falling	226	436	453	946	3224
	Mean		251	384	543	938	3384
[ɤ]	<i>Tơ</i> ‘silk’	Level	276	396	639	1271	3470
	<i>tơ</i> ‘sheet’	Falling	218	483	527	1345	3235
	Mean		247	439	583	1308	3353

Table 2. Vowel characteristics in eight Vietnamese vowels produced by Northern Vietnamese speaker

Vowel	Word	Tone	F0 (Hz)	Duration (ms)	F1 (Hz)	F2 (Hz)	F3 (Hz)
[u]	<i>Tu</i> ‘knock up’	Level	297	400	342	916	3041
	<i>Tù</i> ‘prison’	Falling	219	476	348	839	2698
	Mean		258	438	345	877	2870
[ɯ]	<i>Tư</i> ‘four’	Level	297	414	348	1472	3060
	<i>từ</i> ‘word’	Falling	223	455	387	1368	2892
	Mean		260	435	367	1420	2976
[o]	<i>Tô</i> ‘bowl’	Level	300	408	582	992	3054
	<i>tô</i> non word	Falling	214	458	393	884	2879
	Mean		257	408	488	938	2967
[ɤ]	<i>Tơ</i> ‘silk’	Level	290	413	567	1400	3254
	<i>tơ</i> ‘sheet’	Falling	207	452	555	1310	3176
	Mean		249	432	561	1355	3215

The mean F1 and F2 of Vietnamese Northern and Central vowels are plotted in Figure 3. The vowels [u], [ɤ] and [o] produced by two speakers were quite similar in the vowel space. The vowel that was the

most different from the two speakers was [u]. However, this vowel was different in an unexpected way: Kirby (2011) claimed that Northern [u] was mid-centralized. Yet in this study, the Central speaker produced more mid-centralized [u] than the Northern speaker. Shport (2019) found that Central Vietnamese [u] was partly assimilated to English [ʊ] (32%), [ʌ] (26%) and [u] (22%). Thus, if acoustic similarity between the native vs. non-native vowels could predict the assimilation patterns as PAM claims, Central Vietnamese [u] in this study may be perceived more as English lower [ʊ] or mid central [ʌ] than high back [u] because it was more mid-centralized.

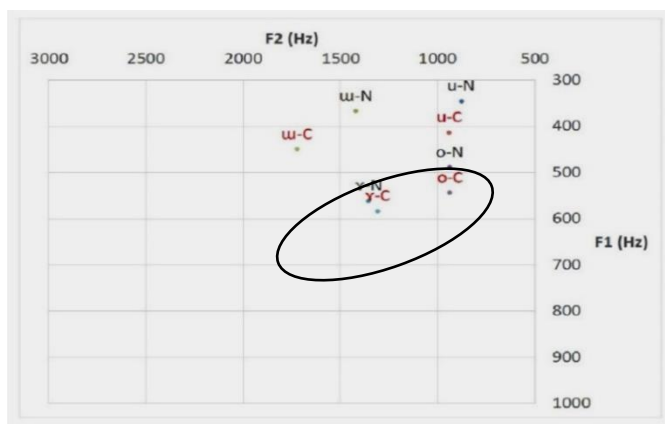


Figure 3. Mean F1 and F2 values of Northern and Central Vietnamese speakers

4.2. Assimilation patterns elicited by the speakers in both dialects

A comparison between Vietnamese and English vowels can be found in Figure 4 below. English vowel characteristics were from Hillenbrand, Getty, Clark, and Wheeler (1995). The vowels within the same circles were close in the vowel space.

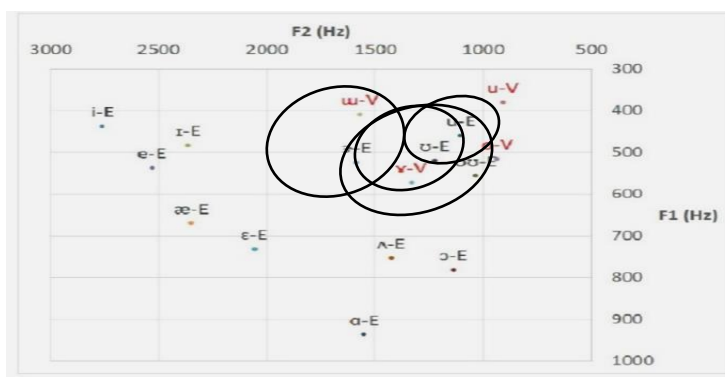
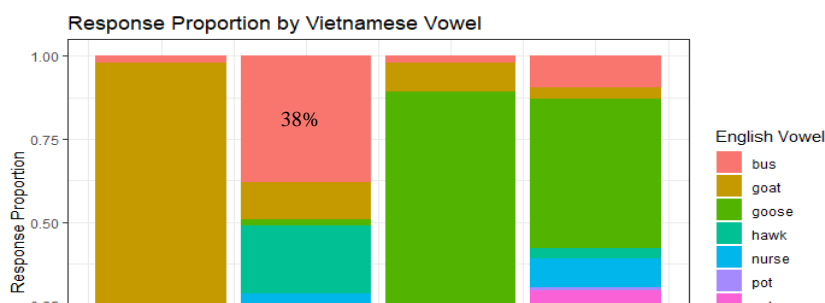


Figure 4. Mean F1 and F2 values of English and Vietnamese vowels

Figure 5 shows the assimilation patterns of the listeners for both Vietnamese dialects.



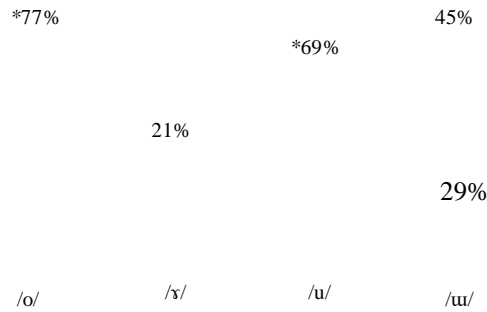


Figure 5. Assimilation patterns of English speakers for both dialects

The results suggested that under the 50% threshold, listeners robustly assimilated Vietnamese [o] to [oo] (77%) and [u] to [u] (69%); [ɤ] and [u̯] were not categorized as any English vowel. Accordingly, the assimilation patterns for [o], [u] was Categorized and [ɤ], [u̯] was Uncategorized. The predicted discrimination accuracy for the Vietnamese vowels from the easiest to the most difficult was: [o-ɤ] = [u-o] = [u̯-u] > [u̯-ɤ] (UC = TC > UU). This was different from Shport (2019): [o-ɤ] = [u̯-ɤ] = [u̯-u] > [u-o]. One notable thing is that even though [ɤ] and [u̯] were uncategorized, they were perceived considerably as [ʌ] (38%) and [u] (45%). Under the above-chance level (14.3%) categorization criterion, these vowels were still assimilated. This indicated a strong tendency of participants' perception towards these vowels. This data will be accounted for by vowel characteristics in 4.3. and 4.4.

4.3. Assimilation patterns by Central and Northern Vietnamese dialects

The assimilation patterns for both dialects were also found for Central Vietnamese dialect. The specific proportion for this dialect can be found in Figure 6.

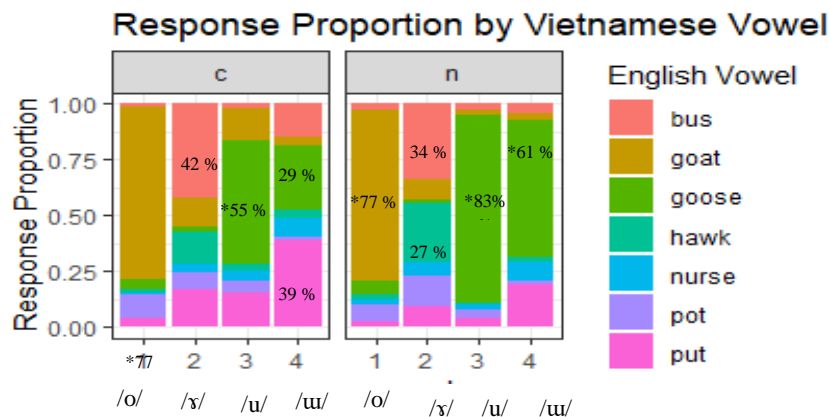


Figure 6. Assimilation patterns of English speakers by dialect

The results suggested that listeners had different assimilation patterns for Northern Vietnamese dialect compared with overall and Central Vietnamese dialect. The first difference was in the vowel [u̯]. Central Vietnamese [u̯] was more mid-centralized than Northern Vietnamese [u̯] and thus Central Vietnamese [u̯] was expected to be perceived more as [ʊ] or [ʌ] than [u]. This effect was indeed observed clearly: Central Vietnamese [u̯] was perceived the most as [ʊ] (39%). This proportion for [u] was lower: 29%. For Northern Vietnamese [u̯] as less mid-centralized, this vowel was assimilated to English [u]: 61%

and the proportion for [o] and [ʌ] was very low: 19% and 4% respectively. This suggested that listeners tended to perceive Central Vietnamese [u] as the lower [o] and mid central [ʌ] than [u] and the opposite was found for Northern Vietnamese [u]. The second notable effect was the difference in the Vietnamese [u]. Northern Vietnamese [u] was higher than Central Vietnamese [u] and this gap was quite big compared with other vowels. This resulted differences in listeners' perception: although Vietnamese [u] was categorized as English [u] in both dialects, the percentage for Central Vietnamese was only 55%, very low compared to Northern Vietnamese: 83%. This indicated that listeners robustly mapped Northern Vietnamese [u] to English [u] using vowel height: the higher the vowel, the more it was perceived as English [u]. Since English [u] has been more fronted, vowel backness may not serve as a reliable cue for the listeners in this case.

5. Implications and conclusion:

This study aimed to find out how English speakers from the Midwest U.S. perceived Vietnamese vowels contrasting in rounding [o ɤ u u] based on PAM framework. The experiment was a replication of Shport (2019) vowel categorization experiment with the English speakers from a different dialect and stimuli from Central and Vietnamese dialect. Overall, the results suggested the followings:

First, the assimilation patterns of participants in general was [o] to [oo] (77%), [u] to [u] (69%); [ɤ] and [u] were not categorized as any English vowel. Accordingly, the assimilation patterns for [o], [u] was Categorized and [ɤ], [u] was Uncategorized. The predicted discrimination accuracy for the Vietnamese vowels from the easiest to the most difficult is: [o-ɤ] = [u-o] = [u-u] > [u-ɤ] (UC = TC > UU). This was different from Shport (2019): [o-ɤ] = [u-ɤ] = [u-u] > [u-o]. This suggested a difference between English speakers from Southern U.S. in Shport (2019) and Midwest U.S. in the current experiment in which cue they used to map between the native and non-native vowels. Specifically, Southern U.S. speakers prioritized vowel rounding to map between Vietnamese [u] to English [oo]. Midwest U.S. speakers prioritized vowel height to map between Vietnamese [u] to English [u]. Southern Vowel Shift may play a role in this difference. Future studies are needed to examine further into this issue. This difference is summarized in Table 3.

Table 3. Assimilation patterns between Southern U.S. and Midwest U.S. speakers

	Southern U.S. speakers (Shport, 2019)	Midwest U.S. speakers (Current study)
[o]	[oo]	[oo]
[ɤ]	[ʌ]	Uncategorized
[u]	[oo]	[u]
[u]	Uncategorized	Uncategorized

Second, Central Vietnamese [u] was more mid-centralized than Northern Vietnamese [u] and this was unpredicted based on Kirby (2011). This resulted in a strong dialect effect in participants' response (Central Vietnamese [u] was perceived the most as [o] and Northern Vietnamese [u] was assimilated to [u]) and suggested that acoustic similarity between native and non-native vowels could predict the assimilation patterns very accurately in this case. Another dialect effect to be found was the difference in [u]. Although both Central and Northern Vietnamese [u] was categorized as English [u], the proportion for Northern Vietnamese [u] was very high compared with Central Vietnamese. This was another indication that English listeners prioritized vowel height to map between non-native and native vowels because Northern Vietnamese [u] was higher than Central Vietnamese [u].

Third, the study demonstrated that language students may perceive ‘similar’ L2 phones as their L1 phones (e.g. the short English [ɪ] may be assimilated to Vietnamese long [i:]). Given the aforementioned findings, the study suggests the following implications for language teachers:

First, teachers should be aware of this outcome and help their students distinguish the very subtle differences between these ‘similar’ sounds. They can use different teaching techniques to help them perceive them correctly (e.g. listen to minimal pairs, listen and guess the sounds, listen and check the pronunciation in dictionary, etc.). Teachers need to emphasize the differences and ensure that they can perceive the sounds correctly through practical techniques, not only through theoretical instructions. Indeed, it is the first step for correct production.

Second, to help students produce correctly similar sounds, teachers can ask students to do different speaking tasks such as reading out loud, recording their own voice, practicing minimal pairs, reading after the recording, etc. Teachers need to keep in mind that correct perception comes before correct production and so their teaching techniques should follow this sequence. Lastly, teachers need to be patient if learners keep substituting L2 sounds with ‘similar’ L1 sounds because it is a natural phenomenon. To perceive/produce these sounds correctly, students need to devote lots of efforts into the process. Also, teachers should not insist that students need to produce L2 sounds perfectly at the beginning. Teachers should instruct them to realize their mistakes over time by comparing their production with model native speakers.

References

- Best, C.T. (1995). A direct realist review of cross-language speech perception. In W. Strange (Ed). Speech perception and linguistic experience. *Issues in cross-language research* (pp.171-204). Timonium, MD: York Press.
- Clopper, Cynthia G., Pisoni, D.B., & Kenneth, D.J. (2005). Acoustic Characteristics of the vowel systems of six regional varieties of American English. *The Journal of the Acoustical Society of America*, 118(3), 1661-1176.
- Erika, S.L. (2009). On the assimilation-discrimination relationship in American English adults’ French vowel learning. *The Journal of the Acoustical Society of America*, 126(5), 2670-2682. <http://doi.org/10.1121/1.3224715>
- Hillenbrand, J., Getty, L.A., Clark M.J., and Wheeler, K. (1995). Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America*, 97(5), 3099-3111.
- Jacewicz, E., Fox., R.A., & Salmons, J. (2011). Cross-generational vowel change in American. English. *Language Variation and Change*, 23(11), 45-86. <http://doi.org/10.1017/S0954394510000219>.
- Kirby, J.P. (2011). Illustration of the IPA: Vietnamese (Hanoi Vietnamese). *Journal of the International Phonetic Association*, 11(3), 381-392.
- Labov, W., Ash, S., & Boberg, C. (2006). *Atlas of North American English: Phonetics, phonology, and sound change*. New York: Mouton de Gruyter.
- Pham, H.T. (2002). Vietnamese tone: Tone is not pitch. (Doctoral dissertation). Retrieved from <http://www2.lib.ku.edu/login?url=https://search-proquest-com.www2.lib.ku.edu/docview/85569726?accountid=14556>.
- Shport, I.A. (2019). Perception of Vietnamese back vowels contrasting in rounding by English listeners. *Journal of Phonetics*, 73, 8-23.
- Simon, E., Debaene, M., & Van Herreweghe, M. (2015). The effect of L1 regional variation on the perception and production of standard L1 and L2 vowels. *Folia Linguistica*, 49(2), 521–553.

CÁCH NGƯỜI MỸ TIẾP NHẬN NGUYÊN ÂM TIẾNG VIỆT PHÂN BIỆT BỞI ĐỘ TRÒN MÔI

Tóm tắt: Nghiên cứu sau đây tìm hiểu cách người nghe đến từ Trung Mỹ tiếp nhận các nguyên âm được phân biệt bởi độ tròn môi trong tiếng Việt. Trong nghiên cứu, người phát âm nói tiếng Trung và Bắc

Việt Nam. Kết quả cho thấy các âm [o] [u] tiếng Việt đã được người nghe đồng hóa với các âm [oo] (77%) [u] (69%). Các âm [ɜ] [ʊ] không bị đồng hóa với nguyên âm tiếng Anh nào. Kết quả gợi ý khi nghe các âm ngoại lai, người nghe từ Trung Mỹ có xu hướng dùng độ cao để so sánh với âm trong tiếng mẹ đẻ. Tuy nhiên người nghe từ Nam Mỹ trong nghiên cứu của Shport (2019) lại dùng độ tròn môi. Ảnh hưởng của phương ngữ cũng được ghi nhận khi âm [u] trong tiếng Bắc được đồng hóa sang [u] (61%) nhưng âm [u] trong tiếng Nam lại được tiếp nhận nhiều nhất là [ʊ] (39%). Vì vậy, sự tiếp nhận âm ngoại lai có thể được ảnh hưởng bởi cả phương ngữ của người nghe và người nói. Nghiên cứu cho thấy giáo viên tiếng Anh nên nhận thức được điều này để có thể dùng những chiến lược thích hợp cho lớp học.

Từ khóa: Mô hình tiếp nhận đồng hóa, đặc thù ngôn ngữ hai, tiếp nhận ngôn ngữ hai, nguyên âm trong tiếng Việt