Journal of Second Language Teaching and Research. Volume 7

Adult L2 Japanese learners' production and perception of Vietnamese monophthong vowels

Đích Mục Đào, University of Social Sciences and Humanities, Vietnam National University-Ho Chi Minh City

Anh-Thư T. Nguyễn, Mountain Creek

Abstract

This paper reports a study that investigated the role of prior native or first language (L1) phonological and phonetic learning on the integration of vowel quality features in the acquisition of second language (L2) vowels by examining adult L2 Japanese learners' perception and production of Vietnamese monophthong vowels in an identification, an imitation and a read aloud tasks. Two groups of participants took part in the study (11 control Vietnamese, 10 Japanese learners of Vietnamese). The stimuli consisted of 9 Vietnamese monophthongs /i, e, ɛ, a, ɔ, o, u, ɤ, ɯ/ and 5 simple Japanese vowels /i, e, a, o, ɯ/. The results showed that Japanese learners of Vietnamese failed to distinguish the Vietnamese vowel pairs /ɛ/-/e/, /o/-/ɔ/ and /u/-/ɯ/ accurately in their perception. In terms of production, Vietnamese vowels /ɛ/ and /e/ merged in vowel space. Moreover, the three Vietnamese vowels /ɔ/, /o/ and/ɤ/ produced by Japanese learners in both production tasks tend to cluster together. Vietnamese vowels /u/ and /ɯ/ produced by Japanese learners also overlapped in vowel space. In general, the findings of this study showed that Japanese learners transfer their L1 vowel quality features into the production of Vietnamese vowels.

Keywords: *vowels; Japanese; Vietnamese; second language acquisition; acoustic phonetics*

Introduction

Learning the phonological system and its phonetic realisations in a second language (L2) is a challenge that learners face in their learning process. In order to account for difficulty that second language learners may face, several models have been proposed that describe the process of non-native speech acquisition. The perceptual assimilation model (PAM; Best, 1995) and speech learning model (SLM; Flege, 1995) are the two major models.

These influential models of speech acquisition both grant an important role to prior linguistic experience of learners, rather than to a universal tendency observed with learners of different first languages. Both models assume that naive second language learners evaluate L2 sounds using their L1 system (assimilation). The assimilation of an L2 sound depends on two factors: (a) whether there exists an equivalent L1 category and (b) its associated

goodness of fit (ranging from marginal to perfect) in the equivalent L1 category. The accuracy of discrimination between contrasting L2 sounds depends on the similarity of their assimilation patterns into L1 categories.

According to PAM (Best, 1995), a given non-native phone may be perceptually assimilated to the native system of phonemes in one of three ways: (1) as a Categorized exemplar of some native phoneme, for which its goodness of fit may range from excellent to poor; (2) as an Uncategorized consonant or vowel that falls somewhere in between native phonemes (i.e., is roughly similar to two or more phonemes); or (3) as a Non-assimilable nonspeech sound that bears no detectable similarity to any native phonemes. Adults' discrimination of a non-native contrast is predicted to depend on how each of the contrasting phones is assimilated. Several pairwise assimilation types are possible. The non-native phones may be phonetically similar to two different native phonemes and assimilate separately to them, which was termed Two Category assimilation. Both may, instead, assimilate equally well or poorly to a single native phoneme, termed Single Category assimilation. Or both might assimilate to a single native phoneme, but one may fit better than the other, termed a Category Goodness difference. Alternatively, one non-native phone may be Uncategorized, as defined above, while the other is Categorized, forming an Uncategorized–Categorized pair. Or both non-native phones might be Uncategorized speech segments. Finally, the two phones' articulatory properties may both be quite discrepant from any native phonemes, and be perceived as Non- Assimilable nonspeech sounds.

Flege (1995) proposed the notion of a *perceptual equivalence* class to account for the purported effect that some non-native vowels are more readily accommodated than others by second language learners. Certain L2 sounds are sufficiently phonetically different from their nearest L1 targets to be perceived as ‘‘new’’ or ‘‘foreign’’, whereas others are sufficiently close to L1 targets to be classified as ‘‘similar’’, though not identical to some L1 phonemic target. Flege (1987) was not very explicit about the phonetic criteria which determine an L2 vowel's classification as ‘‘similar to an L1 target’’ or as ‘‘new’’, but subsequently (Flege & Munro, 1994) assumed that phonetic distance between vowels could be related directly to distances between point targets in the Bark-scaled F 1 -f 0 / F 2 -F 1 plane (Syrdal & Gopal , 1986).Flege (1995) also pointed to the importance of the relationship between perception and production. He hypothesized that accurate perception of L2 sounds will eventually lead to the successful production of L2 phones. If an L2 learner shows difficulty in discriminating L2 contrasts, the learner would also have the same difficulty producing correct L2 phones in L2 learning.

Studies on Japanese acquisition of L2 vowel system are still rare. Ingram and Park (1997) examined the perception and production of Australian English monophthongal non-back vowels: (/i, ɪ, e, æ, a:/) by native speakers of Korean and Japanese, at two levels of English language experience. They also examined prototypicality ratings, or perceived similarities of the foreign vowels to their nearest native (L1) phonemic targets, to evaluate models of cross-language vowel perception. The effects of L1 phonological learning on vowel perception were observed in the tendency of the Japanese, but not the Korean listeners, to normalize tokens of non-native vowels for speaker-dependent durational variation, consistent with the respective phonological roles of vowel length in Japanese and Korean.

A number of other experimental studies have investigated the behaviour of native speakers of Japanese learning English vowels (Tsukada, 1999, 2009; Lambacher, Martens, Kakehi, Marasinghe& Molholt, 2005; Strange et al., 1998). Generally, Japanese speakers show difficulty learning some English vowels. Lambacher et al.(2005) showed that Japanese speakers had difficulty identifying back vowels /ʌ/ and /ɔ/, and open vowels /æ/ and /ɑ/ in American English. Strange et al. (1998) concluded that the “point vowels” (/i/ /ɑ/ /u/) of American English were considered as most similar to the corresponding Japanese vowels (/i/ /a/ /u/) in a relatively stable manner, but less so as long as other vowels are concerned.Tsukada (1999) reported that the English vowels produced by Japanese learners revealed extensive overlap between neighbouring vowels, especially for English /ɑ/-/ʌ/ and /ʊ/-/u/. In other words, Japanese learners seem to assimilate English /ɑ/-/ʌ/ and /ʊ/-/u/ into the single categories of Japanese /a/ and /u/ in production, respectively. Japanese learners seem to use a vowel category that encompassed Japanese /a/, English /ɑ/ and English /ʌ/ and another that encompassed Japanese /u/, English /u/ and English /ʊ/. As a result, the use of the F1 and F2 dimensions for each English vowel was much more widely dispersed in English production by Japanese learners than in production by English monolinguals. Kamiyama and Vaissiere (2009) conducted a series of three experiments to examine the behaviour of Japanese-speaking learners of French (JSL) concerning the perception and production of French /u/, /y/ and /ø/, compared with English-speaking learners of French, in order to examine how L2 vowels which do not have phonemic and/or phonetic equivalents in L1 are acquired by learners. The results indicate that JSL tend to produce French /u/ with a high F2 (> 1000 Hz), which is heard as /ø/ by native French (NF) listeners. They suggest that French /u/ is considered by JSL as phonemically similar to Japanese /u/ (as a high back vowel) and produced as such, while the phonetic realisation of French /u/ is new and

different; and that /y/ is considered as both phonemically and phonetically new, and /ø/ as phonemically new but phonetically(acoustically) similar. They concluded that phonemically similar but phonetically new vowels seem to be most difficult to learn to produce accurately for foreign/second language learners.

Studies on the acquisition of Vietnamese vowels as an L2 are even rarer. Winn et al. (2008) investigated Vietnamese monophthong vowel production by native and American adult learners. Their results suggest that American adult learners struggled to produce the opposition between the central /ɯ/ and back vowel /u/. The learners showed an insufficient advancement separation of these vowels as compared to native speakers. In a recent study, Đào and Nguyễn (2018)examined adult L2 Korean learners' production of Vietnamese monophthong vowels in an imitation and a read aloud tasks. Three groups of participants took part in the study (11 control Vietnamese, 11 Korean learners of Vietnamese, and 10 control Korean). The stimuli consisted of 9 Vietnamese monophthongs /i, e, ɛ, a, ɔ, o, u, ɤ, ɯ/ and 8 Standard Korean vowels /i, ɛ, e, a, o, u,ʌ, ɨ/. The results showed that Vietnamese vowels /ɛ/ and /e/ produced by Korean learners merged in vowel space, proving how a phonemic merger in L1 can influence speakers' perception and production of non-native vowels. Moreover, the three Vietnamese vowels /ɔ/, /o/ and/ɤ/ produced by Korean learners in both tasks tend to cluster together. In general, the findings of their study showed that Korean learners transfer their L1 vowel quality features into the production of Vietnamese vowels.

In this paper, we present new data and results on the role of prior L1 phonological and phonetic learning on the integration of vowel quality features in the production of L2 vowels. Specifically, this study investigates adult L2 Japanese learners' production and perception of Vietnamese monophthong vowels. The findings of this study will have an original and significant contribution to the literature because first, it presents a novel comparison: the acquisition of Vietnamese as an L2 is still understudied. Second, it contributes to the understanding of the process and nature of second language acquisition and has implications for teaching of Vietnamese as a second or foreign language.

Vietnamese and Japanese vowels

The Vietnamese vowel system contains 9 long vowels, 2 short vowels and 3 diphthongs. The long vowels are /i, ɛ, e, a, ɤ, o,ɔ,u, ɯ/; short vowels are /ɐ ʌ/; diphthongs are /ie, ɯɤ, uo/ (Đinh & Nguyễn, 1998). The 9-vowel system analysis (Nguyễn, 1949, 1959; Haudricourt, 1952; Đoàn, 1999; Kirby, 2011) lists these vowels /i e ɛ a ɔ o u ɤ ɯ/ as phonemes and the

vowels /ɐ ʌ/ as allophones of phonemes /a ɤ/. These 9 vowels can occur in both open and closed syllables, while the vowels /ɐ ʌ/ only occur in closed syllables.

Vietnamese is a tonal language which uses pitch to distinguish lexical meaning. The standard Northern dialect has six lexical tones: *level, falling, curve, broken,rising,* and *dropping* (Đoàn, 1999; Đinh & Nguyễn, 1998; Vũ, 1981). Vietnamese tone is superimposed on monosyllables. The six Northern Vietnamese tones combine complex pitch contours with voice quality distinctions (Brunelle, 2009; Kirby, 2010; Michaud, 2004). Voice quality, particularly the laryngeal features of glottal stop, creakiness and breathiness are distinctive tonal features characterizing Vietnamese tones at the phonetic level. Checked syllables, syllables closed by voiceless stops, bear one of two additional tones, which are sometimes considered allotones of tones Rising and Dropping (they will not be addressed here). In perception, glottalization and direction of contour are the dominant cues in Northern Vietnamese (Brunelle & Jannedy, 2013).

Japanese has two sets of five monophthongal vowels pairs (/i-ii, e-ee, a-aa, o-oo, ɯ - ɯɯ/) which are contrastively short or long (mono-moraic or bi-moraic). Japanese long and short vowels do not differ substantially in quality, in other words, the spectral differences between the five long–short pairs are very small (Hirata & Tsukada, 2009) though the geminate forms are slightly more peripheral than their non-geminatecounterparts (Keating & Huffman, 1984). Hence, Japanese may be said to have aphonological contrast of vowel length, whereas some description of Vietnamese phonology describe ơ/ɤ/ and â/ʌ/ as two vowels with long and short opposition and similarly link a/a/ with ă/ɐ/ (Nguyễn &Macken, 2008; Phạm, 2003). An alternative view asserts that all four vowels are distinct in quality (Thompson, 1959; 1965). In a recent study, Winn et al. (2008)'s data showed quality difference for one pair of vowels in long-short opposition (ơ and â) but not the other (a with ă).It is noted that â /ʌ/and ă /ɐ/ occur only in closed syllables.

In this study, we examined 9 Vietnamese monophthongs /i e ɛ a ɔ o u ɤ ɯ/ in open syllables only. Since Japanese long and short vowels do not differ substantially in quality, we focus on examining five short Japanese vowels, namely /i, e, a, o, ɯ/. The 9 vowels of Vietnamese under investigation in terms of tongue raising and advancement are shown in Table 1, each phonetic symbol is followed by its equivalent letters in parentheses:

Table 1. *The Vietnamese vowels*

Position of the tongue / Openness of the mouth	front	central	back
close	i (i, y)	ɯ (ư)[1]	u (u)
Mid	e (ê)	ɤ (ơ)	o (ô)
open	ɛ (e)	a (a)	ɔ (o)

Study aims and plans

The aim of this study is to investigate the role of prior L1 phonological and phonetic learning on the integration of vowel quality features in the acquisition of L2 vowels by examining adult L2 Japanese learners' perception and production of Vietnamese monophthongs in an identification, an imitation and a read aloud tasks. The study aims to address three research questions:

1) To what extent do Vietnamese and Japanese vowels, as spoken by adult Vietnamese and Japanese, differ or overlap in the acoustic phonetic space?
2) Are the phonetic features of L1 Japanese vowels transferred to L2 Vietnamese and how Japanese learners accommodate to the target Vietnamese vowel systems in both perception and production?
3) How would the amount of similarity between Japanese L1 and Vietnamese L2 sounds determine the degree of L1-L2 interaction in late bilinguals?
4) Is there a difference in L2 Japanese speakers' performance in imitation task compared to read aloud task?

Method

Participants

Two groups of participants took part in the study.The control group of 11 Northern Vietnamese (Hanoi) speakers (6 females, 5 males) were international students at Macquarie University and have lived in Australia from 6 months to 1 year. Their average age was 35.3 (standard deviation (SD)=7.2). The Vietnamese subjects can be said to represent Northern Vietnamese (Hanoi) speakers.

[1] The IPA symbol /ɯ/ is used by Đinh & Nguyễn (1998); Thompson (1965); Nguyễn (1997); Đoàn (1977); Kirby (2011), and Emerich (2012) while some other authors (Han, 1966; Winn et al, 2008; and Brunelle, 2015)use the symbol/ɨ/ for this vowel.

The L2 learners of Vietnamese consisted of 10 native Japanese speakers (6 females, 4 males) recruited from students/learners of Department of Vietnamese Studies, University of Social Sciences and Humanities, Vietnam National University Ho Chi Minh City. They all came from Tokyo and have lived in Vietnam for 6 months to 1 year.Their average age was 23 years old (SD=1.7) and their average length of learning Vietnamese was more than one year (mean=14.3 months). In the intermediate Vietnamese language courses, Japanese or other foreign learners basically learn subjects such as Vietnamese vocabulary, Vietnamese grammar, Vietnamese culture, communication skills (listening comprehension in daily situation, radio and television), conversation, reading (newspapers and formal documents), and writing (informal and formal). Because the Japanese leaners started learning Vietnamese at the average age of 19.5 years, they can be considered as late learners of L2. Also, since they were learning in the intermediate Vietnamese language courses, their level of Vietnamese can be considered as intermediate level.

Stimuli

The experiment used open syllables with the initial stop consonant /t-/ and the nine Vietnamese vowels /i/, /e/, /ɛ/, /ɯ/, /ɤ/, /a/, /u/, /o/, /ɔ /. These vowels were then embedded in /t_/ carrier words. Each word independently carried one of the six Northern Vietnamese tones (see Table 2). The total number of items included:9 simple vowels x 6 tones, totalling 54 items. The syllables used in the study are all "legal" syllables, most of which were familiar to the participants.

Table 2. *Vietnamese vowel stimuli*

Words/IPA symbols	***Level tone***	***Falling tone***	***Curve tone***	***Broken tone***	***Rising tone***	***Dropping tone***
ti /ti/	ti	tì	tỉ	tĩ	tí	tị
tê /te/	tê	tề	tể	tễ	tế	tệ
te /tɛ/	te	tè	tẻ	tẽ	té	tẹ
tư /tɯ/	tư	từ	tử	tữ	tứ	tự
tơ /tɤ/	tơ	tờ	tở	tỡ	tớ	tợ
ta /ta/	ta	tà	tả	tã	tá	tạ
tu /tu/	tu	tù	tủ	tũ	tú	tụ
tô /to/	tô	tồ	tổ	tỗ	tố	tộ
to /tɔ/	to	tò	tỏ	tõ	tó	tọ

Since the Japanese learners learnt Vietnamese with instructors of Northern (Hanoi) dialect and they all spoke Vietnamese with a Northern accent, one male native speaker of Hanoi Vietnamese produced all the stimuli for the Identification and the Imitation task, which were recorded at 44.1kHz using the built-in microphone of a laptop and the Praat software (Boersma & Weenink, 2017). The native speaker of Hanoi dialect was in the age of 40, came from Hanoi and was an instructor of Vietnamese at University of Social Sciences and Humanities, National Vietnamese University of Ho Chi Minh City, Vietnam. The stimuli were randomized in one block with the inter-stimulus interval of 6 seconds. The total duration of the block was 13 minutes. The same stimuli were presented in written form via Powerpoint slides for the Read-Aloud task.

Each of the five Japanese vowels occurs in a C(consonant) -V(vowel)- KI context,namely ch**i**ki(/t**ʃi**ki/), s**e**ki(/s**e**ki/), s**a**ki (/s**a**ki/), k**o**ki (/k**o**ki/) and k**ɯ**ki (/k**ɯ**ki/). The chosen Japanese words were disyllabic because it was impossible to find monosyllabic Japanese words that had the same characteristics as the Vietnamese stimuli. Nevertheless, the difference in syllable length across the word sets in the two languages was not seen to be a limitation in comparing vowel quality in Vietnamese and Japanese since the aim was to examine vowel quality only and not vowel length and this has been proven in previous studies (see Yang, 1996 and Baker & Trofimovich, 2005 for an example of English and Korean vowel comparisons using monosyllabic English and disyllabic Korean word stimuli).

Vietnamese words were presented in Vietnamese orthography and Japanese words were presented in Japanese Rōmaji orthography.
The target words in the read aloud task were elicited in a sentential form in which the target word was embedded in a carrier sentence, with all the carrier sentences having the same grammatical structure. The reason why target words were embedded in the middle of a carrier sentence rather than in citation form is to achieve natural speech and to avoid final lengthening effect.

V + O + imperative particle.

For example,

Đọc lại từ ‘tô’ đi nhé.
Read again word big bow please.
‘Say the word “tô” again, please’

Similarly, the Japanese words were elicited in a sentence form:

Futatabi **"chiki"** to iu

Say *'chiki' again*.

The total number of Vietnamese items analysed included: 9 singleton vowels x 6 tones x 11 speakers, totalling 594 items. The total number of Japanese items consisted of 5 vowels x 10 repeats x 10 speakers x 2 tasks, totalling 1000 items.

Procedures

Identification task

The participants (L2 Japanese learners of Vietnamese) sat individually in a quiet room and listened to the stimuli through a laptop computer. They were provided with an answer sheet with all the stimuli (54 rows with 54 stimuli in total and nine Vietnamese vowels in each row). Then they were instructed to listen to the target Vietnamese word and circle the correct vowels that they hear.

Read-Aloud task

The control Vietnamese participants and L2 Japanese learners of Vietnamese were asked to read aloud the 54 stimuli embedded in the carrier sentences presented on Powerpoint slides (one word for each slide) at their own pace. The order of the stimuli was randomized in a different order from the other task. Their responses were recorded by the Praat program on a laptop computer.

Similarly, the L2 Japanese learners of Vietnamese read the sentences embedding five disyllabic words containing five target Japanese vowels 10 times each which were all later used in the analysis. Their responses were recorded by the Praat program on a laptop computer. It is noted that the same microphone and laptop were used for all recordings.

Imitation task

Only the L2 learners of Vietnamese participated in this task. The participants listened to each stimulus once through headphones and were asked to repeat after it without any visual aid. Their responses were recorded by the Praat program on a laptop computer.

All the participants completed the Read-Aloud task last, while the order of the Identification and Imitation tasks was counterbalanced.

Assessing accuracy in the Read-Aloud and Imitation tasks

The recordings by Japanese speakers were judged by two phonetically trained native speakers of Vietnamese, who further identified the vowel errors made by the participants. The

two native speakers evaluated the recordings and labelled the vowel of each syllable/word with a choice among the nine Vietnamese vowels. When there was any disagreement between them, the item was discarded. The two native judges agreed on most of the tokens (inter-rater agreement was 95% for the Imitation task and 88% for the Read-Aloud task), and their divergence appeared to reflect ambiguity in the productions. The learners' mean percentage accuracy and error rates for the nine Vietnamese vowels in the three tasks were calculated and summarized in confusion matrices, which are provided in Table 4. The final number of vowel judgement used in subsequent analysis were sufficient for reliable statistical analysis.

Data analysis

Acoustic analyses of the Vietnamese and Japanese words from the read aloud tasks by both Japanese and control Vietnamese speakerswere performed to determine: (1) how similar (or different) the nine Vietnamese and five Japanese vowels were across the two languages, and (2) how similar (or different) the Japanese learners' nine Vietnamese vowels were from the target L2 Vietnamese and their L1 Japanese.

Acoustic analyses of Vietnamese and Japanese vowels were limited to the fundamental frequency (*F0*) as well as the first two vowel formants (*F1*, *F2*). Although it is possible that the two languages may differ significantly in other dimensions of vowel acoustics (e.g., vowel duration or diphthongization) and that Japanese learners may exploit these differences to make distinctions across their two languages, analyses of these vowel properties were not possible within the present study because of the differences in syllable length across the word sets in the two languages. More specifically, the vowels were analysed by measuring fundamental frequency (*F0*) as well as the first two vowel formants (*F1*, *F2*) at vowel midpoint. The vowels were measured using the get pitch and formant listing commands from the Praat program (Boersma & Weenink, 2017).

The vowel-formant values (in Hz) were then converted to Bark scale (B) to normalize for gender and age differences in vowel production (Syrdal & Gopal, 1986) by using the formula: $B = 26.81 / (1+(1960 / F)) - 0.53$. Two other measures were derived from the obtained vowel-formant values: B1-B0 (B1 minus B0) and B2-B1 (B2 minus B1). B1-B0 is an estimate of vowel position in the high-low dimension, where lower values represent high vowels and higher values represent low vowels. B2-B1 is an estimate of vowel position in the front-back dimension, where lower values represent back vowels and higher values represent front vowels. The vowels were then plotted in the

acoustic space with B2-B1 values on the X-axis and B1-B0 values on the Y-axis and results were presented in Figures 1-5.

Statistical analysis

First, in order to answer the first research question, cross-language comparisons of vowels produced by the control Vietnamese and Japanese speakers were carried out within each of the vowel sets (Vietnamese /a/-Japanese /a/, Vietnamese /e/-Japanese /e/, Vietnamese /i/-Japanese /i/, Vietnamese/o/-Japanese/o/, and Vietnamese /ɯ/-Japanese/ɯ/. In order to account for the effect of speakers' differences and the intrinsic segmental and tonal effects, a restricted maximum likelihood (REML) applied to mixed model methodology was performed on the vowel height (B1-B0) and vowel frontedness (B2-B1) values. The fixed effect included groups (4 groups: L1 Japanese, Japanese imitation, Japanese read aloud and controlVietnamese). The random effect was speakers. The use of REML overcomes the potentially serious deficiency of the ANOVA-based methods which assumed that data are sampled from a random population and normally distributed. REML also avoids bias arising from maximum likelihood estimators in which all fixed effects are known without errors, consequently tend to downwardly bias estimates of variance components. Moreover, REML can handle unbalanced data. The data analysis was carried out using the SPSS program. The results are reported in Table 3.

Second, in order to answer the second and third research questions, comparisons of vowel pairs, which have potential to overlap or merge, within each speaker group (control Vietnamese, L1 Japanese, Japanese Imitation and Japanese read aloud) were carried out within each of the three vowel sets (/ɛ/-/e/, /ɔ/-/o/-/ɤ/, and /u/-/ɯ/). The fixed effect was vowels. The random effect was speakers. A Tukey post-hoc test was then conducted to determine the significant differences among the levels of the main effects. The results arepresented in appendix 1.

Results

Vietnamese and Japanese cross-language comparisons of vowels.

The mixed effect models showed significant effects for the main factor groups for all vowel pairs (F (3,36) =1.8- 12.6, p<0.05-0.0001). As shown in figure 1 and Table 3, Vietnamese and Japanese vowels differ significantly in terms of vowel height (B1-B0, p<0.02-0.001) for vowels /a/ and/e/. By contrast, there is some overlapping of vowel height between the two languages, specifically between vowels such as Vietnamese /i/ and Japanese/i/: p=0.6 ns, Vietnamese /o/ vs. Japanese /o/: p=0.13 nsand Vietnamese /ɯ/vs. Japanese /ɯ/: p=0.46ns.

In addition, there is also some overlapping of vowel frontedness (B2-B1) between the two languages, particularly Vietnamese /i/ and Japanese/i/: p=0.38 ns., Vietnamese /o/ vs. Japanese /o/: p=0.14 ns., and Vietnamese /ɯ/ vs. Japanese/ɯ/: p=0.87 ns. Particularly, it is shown by figure 1 that Japanese vowel /o/ is shown to be in proximity with Vietnamese /o/ and /ɤ/ (Japanese /o/ vs. Vietnamese /o/: B1-B0: p=0.13 ns., B2-B1: P=0.64 ns). Vowel frontedness (B2-B1) of Japanese /o/ was not different from Vietnamese/ɤ/: p=0.64 ns. and Vietnamese /ɔ/: p=0.14 ns.). Japanese vowel /ɯ/ is also in proximity with Vietnamese vowels /u/ and /ɯ/: B1-B0: Japanese /ɯ/ vs. Vietnamese /ɯ/: p=0.46 ns.; Japanese /ɯ/ vs. Vietnamese /u/: p=0.53 ns.; B2-B1: Japanese /ɯ/ vs. Vietnamese /ɯ/: p=0.87 ns.; Japanese /ɯ/ vs. Vietnamese /u/: p=0.82 ns.

Therefore, it is predicted that Japanese L2 speakers of Vietnamese will have difficulty distinguishing the Vietnamese vowels /o/, /ɔ/ and /ɤ/. They also have problems differentiate Vietnamese vowels /u/ and /ɯ/. By contrast, it is also predicted that they can produce acoustic differences for those L1-L2 vowels pairs that were highly dissimilar, such as /a/and /e/ (p<0.001).

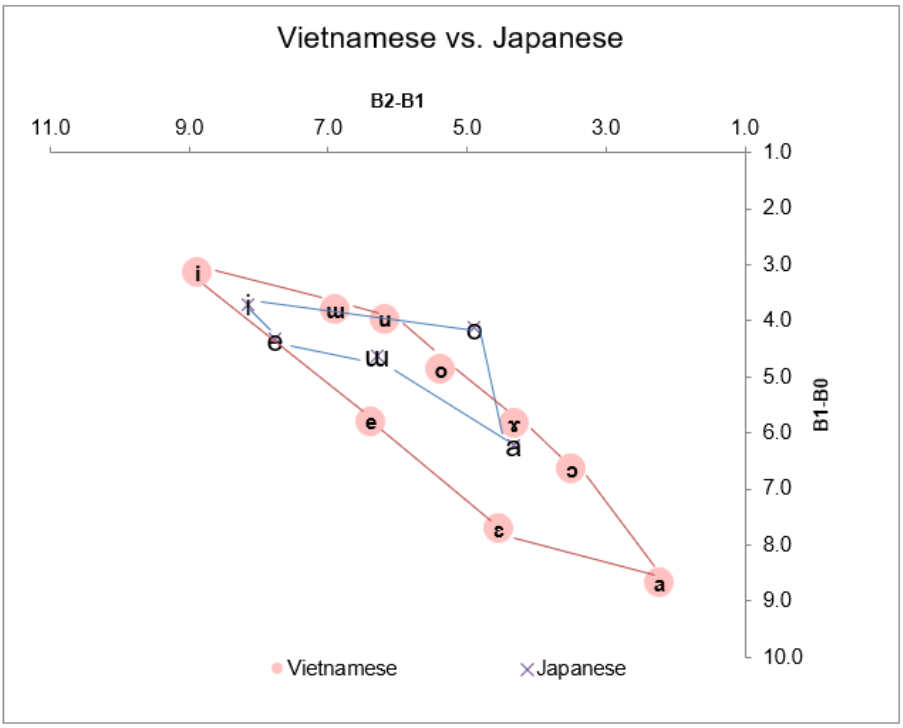

Figure 1. Cross-language comparison of Vietnamese vowels produced by the control Vietnamese and Japanesevowels produced by Japanese learners of Vietnamese. X-axis: B2-B1(Bark): vowel frontedness, y-axis: B1-B0(Bark): vowel height.

Table 3.*Significant levels of the pair-wise comparisons of vowels produced by the four groups: control Vietnamese(V), L1Japanese (J), Japanese imitation (JI), and Japanese read aloud (JR). ns. means non-significance*

	B1-B0					B2-B1				
vowels	V-J	V-JI	V-JR	J-JI	J-JR	V-J	V-JI	V-JR	J-JI	J-JR
a	p<0.0001	p<0.001	p<0.001	p<0.02	p<0.02	p<0.0001	p=0.2 ns.	p=0.37 ns.	p<0.02	p<0.004
ɛ		p<0.05	p<0.02	p=0.7 ns.	p=0.9 ns.		p<0.03	p<0.002	p=0.5 ns.	p=0.7 ns.
e	p<0.02	p=0.06 ns.	p=0.09 ns.			p<0.01	p<0.05	p<0.03		
i	p=0.6 ns.	p=0.28 ns.	p=0.29 ns.	p=0.13 ns.	p=0.13 ns.	p=0.38 ns.	p=0.19 ns.	p=0.27 ns.	p<0.05	p=0.06 ns.
ɔ		p<0.01	p<0.01	p<0.02	p<0.05		p=0.82 ns.	p=0.12 ns.	p=0.22 ns.	p=0.09 ns.
o	P=0.13 ns.	p=0.8 ns.	p=0.7 ns.			p=0.14 ns.	p=0.15 ns.	p=0.47 ns.		
ɤ		p=0.57 ns.	p=0.15 ns.				p=0.75 ns.	p=0.6 ns.		
u		p=0.83 ns.	p=0.69 ns.				p=0.09 ns.	p=0.08 ns.		
ɯ	p=0.46 ns.	p=0.16 ns.	p<0.05	p=0.36 ns.	p=0.28 ns.	p=0.87 ns.	p=0.15 ns.	p=0.47 ns.	p=0.15 ns.	p=0.13 ns.

Vowel error patterns in Identification, Read-Aloud and Imitation tasks

As shown in Table 4, Japanese learners produced significantly more Vietnamese-like vowels in the imitation task than in the read aloud task and they performed worst in the identification task. This is indicated in the result of an ANOVA analysis with a significant effect for tasks ($p<0.0001$). This result suggests that Japanese leaners accommodate to the native Vietnamese speaker to a greater extent in an imitation than in the read aloud task and have difficulties distinguishing Vietnamese vowels in perception.

In the imitation task, there are three main error patterns: Japanese learners have problems distinguishing Vietnamese vowel pairs /o/-/ɔ/ in their production. Japanese learners tend to produce Vietnamese vowel /ɔ/ as /o/ (10%). Similarly, they tend to produce vowel /o/ as /ɤ/ (2%). They also confused between Vietnamese /u/ and /ɯ/ producing /u/ as /ɯ/ (25%). By contrast, they successfully produced vowel /a/, /e/, /i/ and /ɤ/ correctly (100%).

In the read aloud task, they have difficulty producing the vowel /ɛ/ (i.e., /ɛ/ pronounced as /e/ 13%). They also could not distinguish between /o/ and /ɔ/ (/ɔ/ as/o/: 32% and /o/ as /ɔ/: 22%), /o/ and/ɤ/ (/ɤ/ as /o/: 25%), and /u/ and /ɯ/ (/u/ as/ɯ/: 35%). These error patterns support the predictions in the above section.

In the identification task, they confused between /ɛ/ and /e/ perceiving /ɛ/ as /e/ and vice versa (43% and 43% respectively). They also have problems distinguishing /ɔ/, /o/ and /ɤ/ (/ɔ/as /o/:25%, /o/ as /ɔ/:33%, /o/ as/ɤ/: 17%). Surprisingly, they perceived /ɤ/ as /a/: 35%

and /ɤ/ as/ɯ/: 17%. In addition, they confused between /u/ and /ɯ/, perceiving /u/ as /ɯ/: 23% and vice versa /ɯ/ as /u/: 13%.

Table 4. *Vowel error matrix in Imitation, Read-Aloud and Identification tasks. Vertical: Target Vietnamese vowels. Horizontal: Vietnamese vowels pronounced or perceived by Japanese learners. The bolded figures indicate percentage of correct/native like vowel production. The italic figures indicate the major error patterns*.

					Imitation				
Vowels	**a**	**ɛ**	**e**	**i**	**ɔ**	**o**	**ɤ**	**u**	**ɯ**
a	**100**	0	0	0	0	0	0	0	0
ɛ	0	**93**	7	0	0	0	0	0	0
e	0	0	**100**	0	0	0	0	0	0
i	0	0	0	**100**	0	0	0	0	0
ɔ	0	0	0	0	**90**	*10*	0	0	0
o	0	0	0	0	2	**97**	2	0	0
ɤ	0	0	0	0	0	0	**100**	0	0
u	0	0	0	0	0	0	3	**72**	*25*
ɯ	0	0	0	0	0	0	7	2	**92**

					Read aloud				
Vowels	**a**	**ɛ**	**e**	**i**	**ɔ**	**o**	**ɤ**	**u**	**ɯ**
a	**100**	0	0	0	0	0	0	0	0
ɛ	0	**87**	*13*	0	0	0	0	0	0
e	0	2	**98**	0	0	0	0	0	0
i	0	0	3	**97**	0	0	0	0	0
ɔ	0	0	0	0	**60**	*32*	3	0	5
o	0	0	0	0	*22*	**78**	0	0	0
ɤ	0	0	0	0	8	*25*	**67**	0	0
u	0	0	0	0	0	0	3	**62**	*35*
ɯ	0	0	0	0	0	0	3	0	**97**

					Identification				
Vowels	**a**	**ɛ**	**e**	**i**	**ɔ**	**o**	**ɤ**	**u**	ɯ
a	**98**	0	0	0	0	0	2	0	0
ɛ	8	**48**	*43*	0	0	0	0	0	0
e	2	*43*	**53**	0	2	0	0	0	0
i	0	5	2	**93**	0	0	0	0	0
ɔ	7	0	0	0	**43**	*25*	*25*	0	0

o	0	0	0	0	*33*	**50**	17	0	0
ɤ	*35*	0	5	0	3	8	**28**	3	17
u	0	0	0	0	0	0	3	**73**	*23*
ɯ	0	0	0	0	0	2	2	*13*	**83**

Acoustic comparison of Vietnamese vowels produced by Japanese learners and control Vietnamese

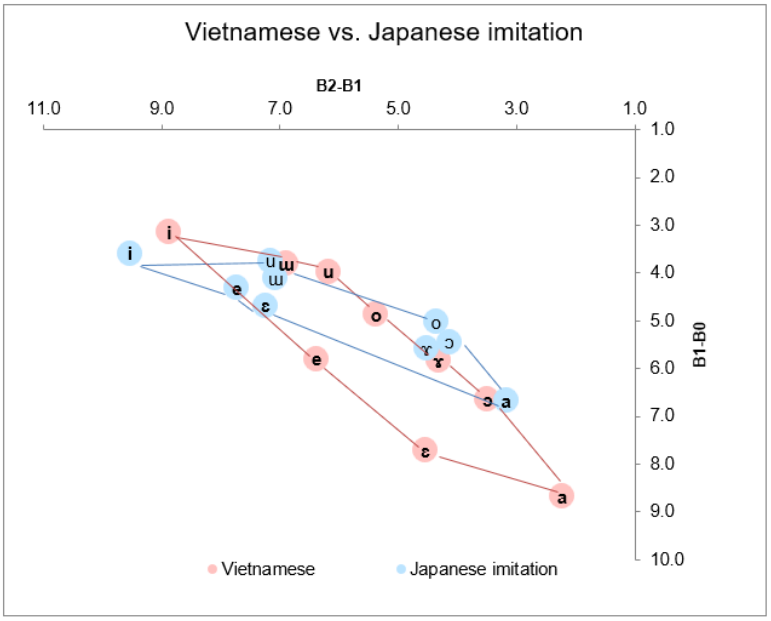

Figure 2. Vietnamese vowels produced by Japanese learners in the imitation task and control Vietnamese. X-axis: B2-B1(Bark): vowel frontedness, y-axis: B1-B0(Bark): vowel height.

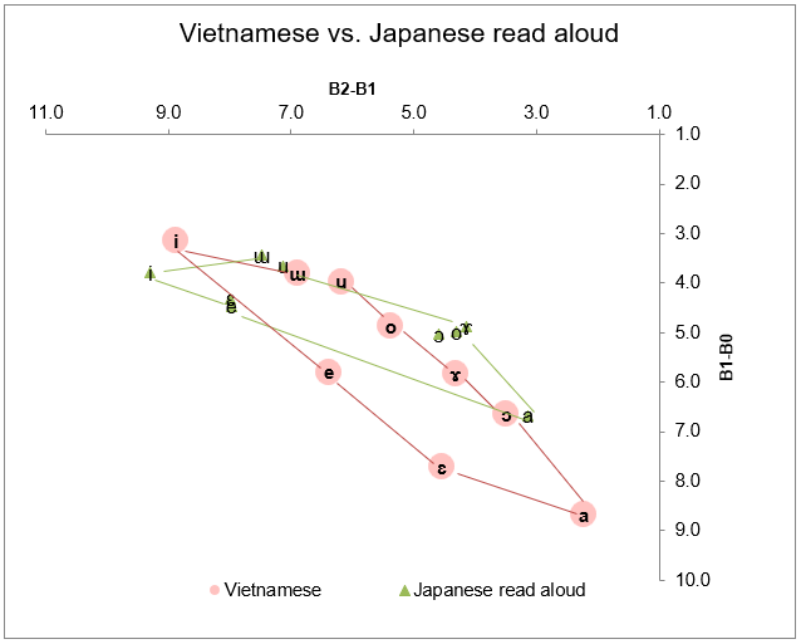

Figure 3. Vietnamese vowels produced by Japanese learners in the read aloud task and control Vietnamese. X-axis: B2-B1(Bark): vowel frontedness, y-axis: B1-B0(Bark): vowel height.

As shown in figures 2 and 3, Japanese learners produced vowels which have significantly higher vowel height than those of control Vietnamese speakers across both tasks: imitation and read aloud. This mirrors the pattern found in section on *Vietnamese and Japanese cross-language comparisons of vowels*: some Japanesevowels (/e/, /o/ and /a/) also have higher vowel height than those of control Vietnamese.

In addition, Vietnamese vowels /ɛ/ and /e/ produced by Japanese learners merged in vowel space, supported by the statistical analysis as reported in appendix 1 and figures 4&5. This is consistent with the identification data and perception results by two phoneticians in section on *Vowel error patterns in Identification, Read-Aloud and Imitation tasks.*

Moreover, the three Vietnamese vowels /ɔ/, /o/ and/ɤ/ produced by Japanese learners in both tasks tend to cluster together. This is also supported by the statistical analysis in appendix 1 and figures 4&5 and consistent with the identification patterns. Specifically, there was no significant difference in vowel height (B1-B0) and the vowel frontedness (B2-B1) of the vowels /o/-/ɔ/-/ɤ/ for both tasks (imitation and read aloud).
The statistical result (appendix 1) indicates that Japanese learners' (JI and JR) Vietnamese vowels /u/-/ɯ/ are of the same vowel height (B1-B0: p=0.13 ns. and p=0.44 ns., respectively). The vowel frontedness of this vowel pair is also the same (B2-B1: p=0.85 ns. and p=0.43 ns., respectively).

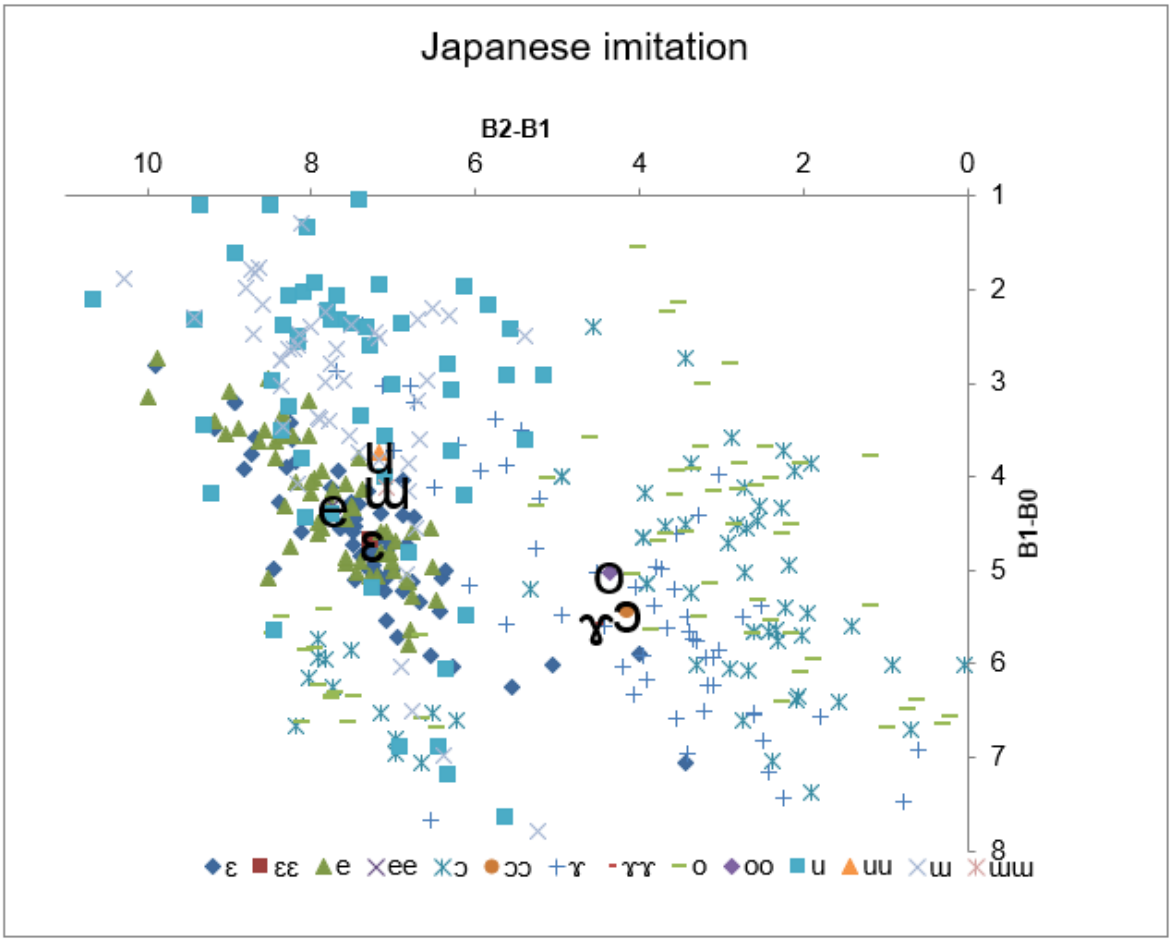

Figure 4. The merging patterns of Vietnamese vowels produced by Japanese learners in the imitation task. X-axis: B2-B1(Bark): vowel frontedness, y-axis: B1-B0(Bark): vowel height. Double IPA symbols in the legend indicate the mean values of the vowels (e.g., ee is the mean value of the vowel e).

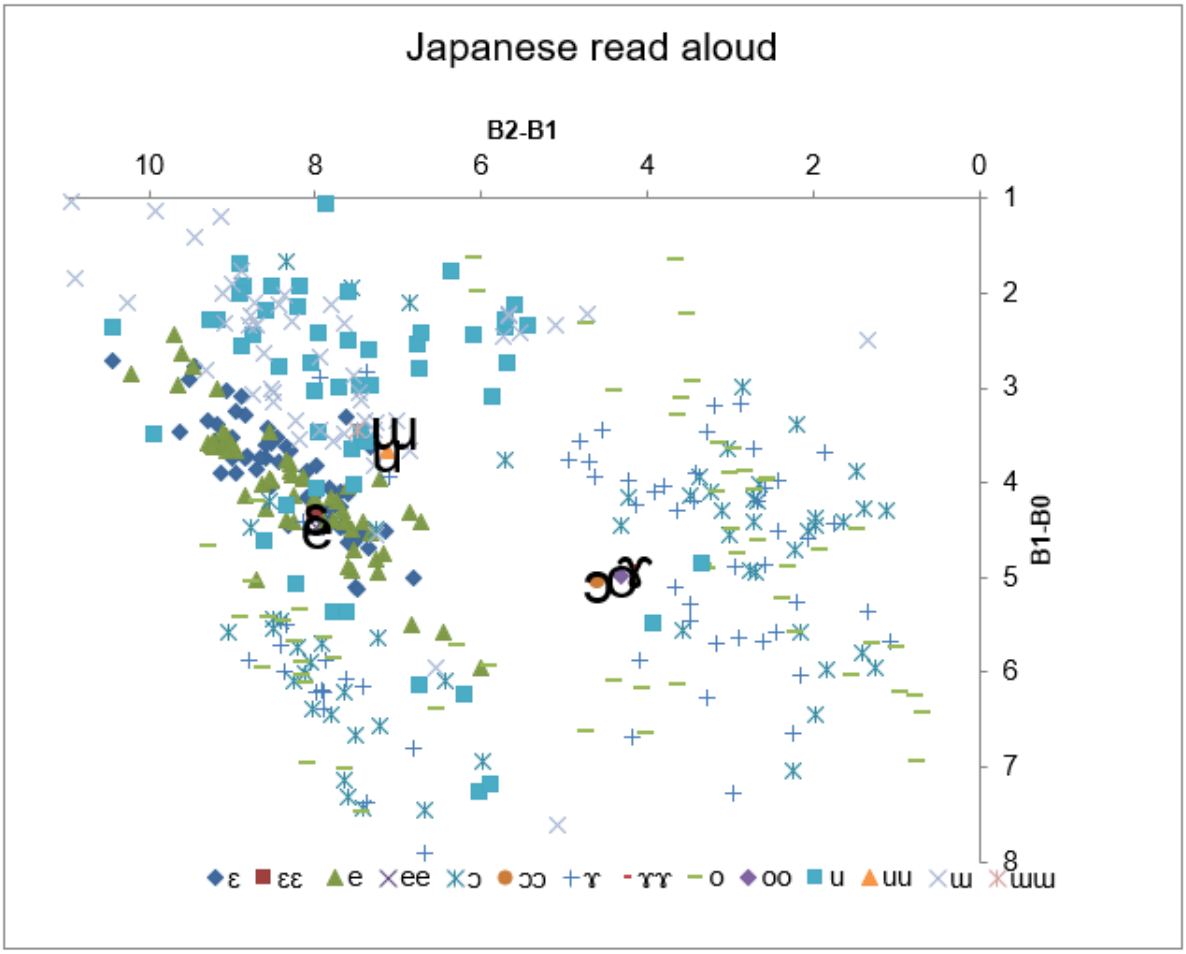

Figure 5. The merging patterns of Vietnamese vowels produced by Japanese learners in the read-aloud task. X-axis: B2-B1(Bark): vowel frontedness, y-axis: B1-B0(Bark): vowel height. Double IPA symbols in the legend indicate the mean values of the vowels (e.g., ee is the mean value of the vowel e).

Discussion

In this section, we summarize and discuss the results by addressing the four research questions raised in the section on *Study aims and plans.*

First, *to what extent do Vietnamese and Japanese vowels, as spoken by adult Vietnamese and Japanese, differ or overlap in the acoustic phonetic space?*

The result on cross language comparison showed that Japanese vowels /a/ and /e/ differ significantly from Vietnamese counterparts /a/ and /e/ in terms of vowel height and vowel frontedness. By contrast, there is some overlapping of vowel height and vowel frontedness between the two languages (particularly Vietnamese /i/ vs. Japanese/i/ and Vietnamese /ɯ/ vs. Japanese /ɯ/). Furthermore, Japanese /o/ is shown to be in proximity with Vietnamese /o/ and /ɤ/ while Japanese /ɯ/ is in immediacy to Vietnamese /u/ and/ɯ/.

Second, *are the phonetic features of L1 Japanese vowels transferred to L2 Vietnamese and how Japanese learners accommodate to the target Vietnamese vowel systems?*
The three Vietnamese vowels /ɔ/, /o/ and/ɤ/ produced by Japanese learners in both tasks tend to cluster together. This may be due to the promixity in acoustic space of the Japanese vowels /o/ to the Vietnamese vowels /ɔ/, /o/ and/ɤ/. This provided insights into how cross-language similarity influenced the L1-L2 interaction. That is, when Vietnamese and Japanese vowels were relatively similar acoustically, the Japanese learners' renditions of L2 (Vietnamese) vowels were strongly "colored" by the acoustic properties of their L1 (Japanese) vowels, consistent with findings by previous studies on Korean learners (Baker & Trofimovich, 2005; Đào and Nguyễn (2018). Furthermore, Vietnamese vowels /ɛ/ and /e/ produced by Japanese learners merged in vowel space. This could be predicted from any of the models of cross-language vowel perception (e.g., Flege, 1995; Best, 1995). It seems most likely that the L2 perception and production differences were expressions of L1 perceptual learning effects. In other words, incorrect perception of L2 phones by adult learners is considered to be due to the assimilation of L2 phones to L1 categories. This can be explained according to Best (1995) model: the Japanese learners may have assimilated Vietnamese vowel /ɛ/ and /e/ contrast to a single Japanese /e/ category since there is no such counterpart of vowel /ɛ/ in Japanese vowel system which shows the case of Single Category type. The same principle applies to the Vietnamese /o/ and /ɔ/ contrast, Japanese has /o/ but not /ɔ/, thus they tend to perceive and pronounce Vietnamese /ɔ/ as /o/, suggesting that they assimilated Vietnamese /o/ and /ɔ/ to a single Japanese /o/. Similarly, Japanese has /ɯ/ but not /u/, as a result, they tend to perceive and pronounce Vietnamese /u/ as /ɯ/. This is consistent with previous results on the English vowels produced by Japanese learners that revealed extensive overlap between neighbouring vowels (Tsukada, 1999).

In addition, the result on L2 vowel production also indicates that some of Japanese learners' Vietnamese vowels are higher than those of the control Vietnamese, suggesting that Japanese leaners tend to transfer their L1 vowel acoustic space into the production of Vietnamese vowels.

Third*, how would the amount of similarity between Japanese L1 and Vietnamese L2 sounds determine the degree of L1-L2 interaction in late bilinguals?*
The results of this study indicated that cross-language similarity indeed influenced how the L2 vowels are produced. That is, vowels that were highly similar across the two languages

were more likely to influence each other (such as /e/, /o/ and/ɯ/) than those vowels that were dissimilar (such as /a/). The late bilinguals in this study produced acoustic differences only for those L1-L2 vowel pairs that were highly dissimilar, such as Vietnamese /a/ and Japanese/a/. Assuming that L1-L2 interaction implies restructuring of the L1 and L2 phonetic system(s), then the degree of acoustic similarity between L1 and L2 sounds constrains what sounds undergo such a restructuring and the degree to which it does so (Trofimovich et al., 2001). The interaction hypothesis (Flege et al., 1995) may explain why cross-language similarity is more likely to determine how adult L2 learners organize their phonetic system(s). Because late bilinguals' L1 categories are fully developed, they are more likely to produce even perceptually dissimilar L1 and L2 sounds with L1-based acoustic properties (Aoyama et al., 2004) and to perceive such L2 sounds in terms of an L1-based category (Guion et al., 2000; Trofimovich et al., 2001). The late bilinguals in this study may require an amount of experience with the L2 that is far greater than that explored in this study (more than one year) in order to overcome the pervasive effect of their L1 on their processing and learning of L2 sounds (Flege et al., 1995; Trofimovich et al., 2001). These results also suggest that phonemically similar but phonetically new vowelsseem to be most difficult to learn to produce accurately for foreign/second language learners (Kamiyama & Vaissiere, 2009).

Additionally, the result of the study also indicates the effect of Japanese learners' vowel perception on their production. The fact that they failed to distinguish the Vietnamese vowel pairs /ɛ/-/e/, /o/-/ɔ/ and /u/ as /ɯ/ accurately in their production (in imitation and read aloud tasks) is strongly supported by their perception patterns in the identification task, suggesting that they have problems perceiving the vowel contrast in L2. Flege (1995) pointed to the importance of the relationship between perception and production. He hypothesized that accurate perception of L2 sounds will eventually lead to the successful production of L2 phones. If an L2 learner shows difficulty in discriminating L2 contrasts, the learner would also have the same difficulty producing correct L2 phones in L2 learning. Further research examining how cross-language similarity influences the perception of the L2 may indicate to what extent this ability constrains both the perception and the production abilities of bilinguals.

Fourth*, is there a difference in L2 Japanese speakers' performance in imitation task compared to read aloud task?*

The results showed that Japanese learners produced significantly more Vietnamese-like vowels in the imitation task than in the read aloud task. This result suggests that Japanese leaners accommodate to the native Vietnamese speaker to a greater extent in an imitation

than in the read aloud task. This can be explained by the phonetic convergence effect which is defined as the process by which a talker takes on acoustic characteristics of the individual that he or she is interacting with (Babel, 2012). The results revealed a significant convergence with the model in the task in which speakers were required to immediately repeat after the model voice (imitation task) compared to the task in which they read orthographic representations of the words (read aloud task). Hence, it suggests that foreign language learners are able to modify their productions of non-native vowels as a result of exposure to the model. The result that Japanese learners become more similar in their production to the target language speaker in the imitation task would also imply that fine-grained phonetic details are not filtered out in speech perception and detailed auditory traces associated with perceived words are stored in memory and are then used for production (Dufour & Nguyen, 2013).

Conclusion

In summary, this paper presents new data and results on the roles of language-specific phonological learning and inherent phonetic contrastiveness in the production of non-native vowels. The results of this study provide evidence that Japanese learners transfer their L1 vowel quality features into the production and perception of Vietnamese vowels. The findings of this study have an original and significant contribution to the literature because first, it presents a novel comparison: the acquisition of Vietnamese as an L2 is still understudied. Second, it contributes to the understanding of the process and nature of second language acquisition and has implications for teaching of Vietnamese as a second or foreign language. The difficulties experienced by Japanese learners as found in this study may guide future research in examining teaching strategies. While the results of this study are specific to Japanese and Vietnamese, their implications can be extended to the acquisition of other languages. Somelimitations of this study include a limited number of participants, inclusion of only one speaker for the stimuli for the identification and imitation tasks and analysis of F0, F1 and F2 only.

Acknowledgements

We would like to thank the subjects for their voluntary participation in the experiment. We also thank the anonymous reviewers for their constructive comments.

Biodata

Đích Mục Đào completed his Ph.D. in Linguistics 2013 at the University of Queensland, Australia. He did his MA in Linguistics 2002 and BA in Literature & Linguistics 1991 at University of Social Sciences and Humanities, Vietnam National University, Ho Chi Minh City. His research interests include acoustic phonetics, Vietnamese studies and teaching the Vietnamese language for foreigners. Email:dichvns@hcmussh.edu.vn

Anh-Thư Thị Nguyễn completed her MA in Applied Linguistics 1999, Ph.D in Linguistics 2003 and Post-Doctoral research at the University of Queensland, Australia. She did her BA in TESOL at Can Tho University , Viet Nam 1994. Her research interests include Vietnamese prosody, Second language phonetics and phonology, Vietnamese L1 attrition and acoustic phonetics.

References

Aoyama, K., Flege, J. E., Guion, S. G., Akahane-Yamada, R.,& Yamada, T. (2004). Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese /r/ and English /r/ and / l/. *J. Phonetics,23*, 233-250.

Babel, M. (2012). Evidence for phonetic and social selectivity in spontaneous phonetic imitation. *Journal of Phonetics*, *40*, 177-189.

Baker, W., & Trofimovich, P. (2005). Interaction of native- and second-language vowel system(s) in early and late bilinguals. *Language & Speech, 48*, 1–27.

Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: issues in cross - language research,* (pp. 171-203). Baltimore: York Press.

Boersma, P., & Weenink, D. (2017). Praat: doing phonetics by computer (Version 6.0.26). Computer program. Retrieved January 1, 2017, from http://www.praat.org/ California, Los Angeles.

Brunelle, M. (2015). Vietnamese [Tiếng Việt]. In P. Sidwell & M. Jenny (Eds.), The Handbook of Austroasiatic Languages (pp. 909-954). Leiden/Boston: Brill Academic.

Brunelle, M. (2009). Tone perception in Northern and Southern Vietnamese. *Journal of Phonetics, 27*(1), 79-96.

Brunelle, M., & Jannedy, S. (2013). The Cross-dialectal Perception of Vietnamese Tones: Indexicality and Convergence. In D. Hole & E. Löbel (Eds.), *Linguistics of Vietnamese*, Amsterdam: John Benjamins.

Đinh, L. T., &Nguyễn, V. H. (1998). *Cơ cấu ngữ âm tiếng Việt [Structure of Vietnamese phonetics]*. Thành phố Hồ Chí Minh: Giáo Dục Press.

Đoàn, T. T. (1979). *Ngữ âm tiếng Việt[Vietnamese phonetics].* Hà Nội: Đại Học Quốc Gia Hà Nội Press.

Đào, Mục-Đích and Nguyễn, T.A.T. (2018). L1 Korean vocalic transfer in adult L2 Korean learners' production of Vietnamese monophthong vowels. *Asian-Pacific Journal of Second and Foreign Language Education*, 3:13

Dufour, S.,& Nguyen, N. (2013). How much imitation is there in a shadowing task? *Front Psychol*,*4*, 1-7.doi 10.3389/fpsyg.2013.00346

Emerich, G. H. (2012). *The Vietnamese vowel system*. Doctoral dissertation. University of Pensylvania.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233-277). Baltimore: York Press.

Flege, J. E. & Munro, M. J. (1994) Auditory and categorical ef fects on cross-language vowel perception, *Journal of the Acoustical Society of America, 95* (6), 3623 – 3641

Flege, J. E., Munro, M. J., & MacKay, I. R. A. (1995). Factors affecting degree of perceived foreign accent in a second language. *Journal of the Acoustical Society of America, 97*, 3125-3134.

Guion, S., Flege, J., Akahane-Yamada, R., & Pruitt, J. (2000). An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *Journal of the Acoustical Society of America, 107*, 2711-2725.

Han, M. S. (1966). *Studies in the Phonology of Asian Languages 4: Vietnamese vowels*. Los Angeles: Acoustic Phonetics Research Laboratory, University of Southern California.

Haudricourt, A. G. (1952). Les Voyelles brèves du vietnamien. *Bulletin de la Société de Linguistique de Paris, 48*(1), 90-93.

Hirata, Y., & Tsukada, K. (2009). Effects of speaking rate and vowel length on formant frequency displacements in Japanese. *Phonetica, 66*(3), 129-149.

Ingram, J. C., & Park, S-G. (1997). Cross-language vowel perception and production by Japanese and Korean learners of English. *Journal of Phonetics, 25*(3), 343-370.

Kamiyama, T., &Vaissière, J. (2009). Perception and production of French close and close-mid rounded vowels by Japanese-speaking learners. In J. Y. Dommergues, (Ed.), *Revue AILE - LIA, 2*, 9-41.

Keating,P. A.,& Huffman, M. K. (1984). Vowel variation in Japanese. *Phonetica, 41*, 191-207.

Kirby, J. (2010). Dialect experience in Vietnamese tone perception. *Journal of the Acoustical Society of America, 127*(6), 3749-3757.

Kirby, J. P. (2011). Vietnamese (Hanoi Vietnamese). *Journal of the International Phonetic Association,41*(3), 381-392.

Lambacher, S., Martens, W., Kakehi, K., Marasinghe, C., &Molholt, G. (2005). The effects of identification training on the identification and production of English vowels by native speakers of Japanese. *Applied Psycholinguistics, 26*, 227-247.

Michaud, A. (2004). Final consonants and glottalization: New perspectives from Hanoi Vietnamese. *Phonetica, 61*,119-146.

Nguyễn, B. T. (1949). *Chữ và Vần Việt Khoa Học* [Scientific study of Vietnamese letters and syllables]. Sài Gòn: Ngôn Ngữ.

Nguyễn, B. T. (1959). *Ngôn Ngữ học Việt Nam* [Vietnamese linguistics]. Sài Gòn: Ngôn Ngữ.

Nguyễn, Đ.H. (1997). *Tiếng Việt Không Son Phấn* [Vietnamese without veneer]. Amsterdam: John Benjamins.

Nguyễn, T. H.,& Macken, M. (2008). The acquisition of Vietnamese tones: A Study of American Learners. *Studies in Second Language Acquisition, 30*(1), 49-77.

Phạm, H.A. (2003). *Vietnamese tone: A new analysis*. Routledge.

Strange, W., Akane-Yamada, R., Kubo, R., Trent, S. A., Nish, K., Jenkins, J. (1998). Perceptual assimilation of American English vowels by Japanese listeners. *Journal of Phonetics, 26,*311-344.

Syrdal, A. K.,& Gopal, H. S. (1986). A perceptual model of vowel recognition based on the auditory representation of American English vowels. *Journal Acoustical Society of America, 79*, 1086-1100.

Thompson, L.C. (1959). Saigon phonemics. *Language, 35*, 454-476.

Thompson, L.C. (1965). *A Vietnamese reference grammar*. Honolulu: University of Hawai'i Press.

Trofimovich, P.,& Baker, W. (2006). Learning second language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech. *Studies in Second Language Acquisition, 28*, 1-30.

Trofimovich, P., Baker, W., & Mack, M. (2001). Context and experience-based effects on the learning of vowels in a second language. *Studies in the Linguistic Sciences, 31*(2), 167-186.

Tsukada, K. (2009). Durational characteristics of English vowels produced by Japanese and Thai second language learners. *Australian Journal of Linguistics, 29*(2),287-299.

Tsukada, K. (1999). *An acoustic phonetic analysis of Japanese-accented English*. Unpublished Doctoral dissertation. Macquarie University, Sydney, Australia.

Vũ, T. P.(1981).*The acoustic and perceptual nature of tone in Vietnamese* (Unpublished doctoral dissertation). Australian National University, Canberra.

Winn, M., Blodgett, A., Bauman, J., Bowles, A., Charters, L., Rytting, A.,& Shamoo, J. (2008). *Vietnamese monophthong vowel production by native speakers and American adult learners*. Proceedings of Acoustics'08 Paris, pp. 10473-10478.

Yang, B. (1996). A Comparative study of American English and Korean vowels produced by male and female speakers. *Journal of Phonetics, 24*, 245-261.

Appendix 1: Mean and significant levels of pairwise comparisons of vowels

	B1-B0						B2-B1					
	e	ɛ	Sig.				e	ɛ	Sig.			
Japanese	4.33						7.76					
Japanese imitation	4.32	4.68	**p=0.16 ns.**				7.73	7.25	**p=0.14 ns.**			
Japanese read aloud	4.48	4.32	**p=0.55 ns.**				7.97	7.99	**p=0.91 ns.**			
Vietnamese	5.79	6.28	p<0.004				6.38	5.68	p<0.001			
	ɔ	o	ɤ	Sig.			ɔ	o	ɤ	Sig.		
				ɔ-o	ɔ-ɤ	o-ɤ				ɔ-o	ɔ-ɤ	o-ɤ
Japanese	4.12						4.89					
Japanese imitation	5.44	5.03	5.58	**p=0.12 ns.**	**p=0.58 ns.**	**p<0.05**	4.14	4.36	4.53	**p=0.39 ns.**	**p=0.12 ns.**	**p=0.49 ns.**
Japanese read aloud	5.04	4.99	4.89	**p=0.63 ns.**	**p=0.64 ns.**	**p=0.98 ns.**	4.59	4.31	4.15	**p=0.52 ns.**	**p=0.35 ns.**	**p=0.77 ns.**
Vietnamese	6.64	4.87	5.83	p<0.0001	p<0.02	p<0.005	3.48	5.37	4.32	p<0.0001	p<0.05	p<0.01
	u	ɯ	Sig.				u	ɯ	Sig.			
Japanese	4.62						6.3					
Japanese imitation	3.74	4.11	**p=0.13 ns.**				7.16	7.07	**p=0.85 ns.**			
Japanese read aloud	3.66	3.44	**p=0.44 ns.**				7.13	7.48	**p=0.43 ns.**			
Vietnamese	3.96	5.58	p<0.0001				6.18	5.42	p<0.005			