English Vowel Duration Affected By Voicing Contrast In Chinese, Korean, Japanese, And Vietnamese Speakers

Nguyen Van Anh Le, Sophia University, Japan
Mafuyu Kitahara, Sophia University, Japan

The Asian Conference on Language 2020
Official Conference Proceedings

Abstract
Learning a foreign language comes with learning new phonetic contrasts. Adults often have difficulty in learning non-native vowels especially when the vowel inventories in their first and second language are different. This study describes the speech acquisition of English vowel in Chinese, Korean, Japanese, and Vietnamese speakers, and addresses their problem of voicing dependent vowel duration by assessing how vowel quantity is produced before voiced and voiceless consonants. A gating experiment was conducted to assess how reliably speakers can produce English monophthongs and diphthongs. To achieve this overarching purpose, three aims were addressed: (1) to identify the durational properties of English vowels preceding voiced/voiceless consonants of the four languages, (2) to present an overview of the four languages divided in 3 groups: tonal, non-tonal, and pitch accented, (3) to investigate difference in subjects' acquisition of English vowel based on their language systems. Chinese, Vietnamese-L1 English learners (Group 1), whose L1 is tone languages and have no audible released codas, Korean-L1 English learners (Group 2), whose L1 has no vowel quantity distinction (VQD), and Japanese-L1 English learners (Group 3), whose L1 is pitch accented language, has no aspirated codas and has phonemic length contrast (i.e. short and long) VQD. All three groups produced longer vowels before voiced codas with monosyllabic words, however, they did not produce vowel lengthening before English voiced codas at disyllabic words, which was significantly different from English native speakers.

Keywords: vowel duration, voicing contrast, vowel quantity distinction

Introduction

English speech sounds produced by non-native speakers are significantly different from those produced by American English native speakers. It is reported that English vowel duration patterns produced by English-, Chinese-, Korean-, and Japanese native speakers are probably different due to the fact that vowel durations play different roles in these languages. English vowel is considered longer before voiced than voiceless consonant as the duration of stressed vowels in English is primarily governed by the voicing of the following consonant. In adult speech, vowels preceding voiceless consonants are approximately two-thirds the duration of vowels preceding voiced consonants (House, 1961). The aim of this research is to examine the extent to which voicing contrast of word final obstruents transferred to English vowel duration produced by bilingual speakers.

Vowel length as a cue for voicing

In American English, as in most languages, vowels preceding a voiced consonant have longer duration then ones preceding voiceless consonants, it means that speakers of English might use vowel duration distinction as for interpreting English coda voicing distinction in both their production and perception. However, there are several studies that have been examined vowel lengthening before voiced coda produced by Dutch, Arabic, and French speakers of English, and no significant vowel lengthening in their L2 has been observed (Flege & Port, 1981; Mack 1982; Elsendoorn, 1983). These results have first evidenced against vowel lengthening before voiced codas as universal. This study continues at examining the phonemic vowel lengthening patterns of speakers from other L1 backgrounds (i.e Chinese, Vietnamese, Korean, and Japanese) to explore whether vowel duration in L1 determines vowel lengthening patterns in L2. More specifically, if L1 Japanese speakers were found to produce longer vowels before English voiced codas than voiceless ones, it could be supposed that the presence of a phonemic vowel length distinction in speaker's native language can determine their L2 vowel length distinction. In other words, it could be said that the acquisition of L2 vowel patterns depends on L1 vowel system in which the vowel length will be affected. It also may imply that the long and short vowel phonemes in Japanese allow their speakers to acquire long and short L2 allophones. On the other hand, if no vowel lengthening were found on Japanese speakers, the claim that L2 vowel lengthening patterns depends on L1 vowel system could be weakened.

Moreover, it is expected that leaners whose language without vowel quantity distinction as Korean hardly distinguish vowel lengthening before English voiced codas. If L1 Korean speakers were found to effectively produce longer duration of vowels before voiced codas than voiceless ones, it may also weaken the claim that L2 vowel lengthening patterns were affected by L1 vowel system, and vice versa, L1 Korean speakers' inability to produce long and short English allophones could be explained by their inexistence of long and short vowels in their language system.

The purpose of this study is to examine whether L2 speakers of English can interpret English coda voicing distinction in the speech, and the target vowels used in this study are not only monophthongs but also diphthongs to show a closer look on the vowel lengthening patterns of these four major languages in Asia. To achieve this overarching purpose, three aims were addressed:

(1) to identify the duration of English vowels preceding voiced/voiceless codas of four subjects above
(2) to present an overview of the four languages divided in 3 groups: tonal, non-tonal, and pitch accented
(3) to investigate difference in subjects' acquisition of English vowel based on their language systems.

Literature review

This section will give an overview of syllable structures and coda voicing in English, Chinese, Vietnamese, Korean, and Japanese, focusing on aspects of vowel system and vowel quantity distinction. English phonology has a phonemic contrast between voiced and voiceless codas, while Chinese and Vietnamese do not have aspiration in coda position. In addition, neither voiced-voiceless stop contrasts nor long-short vowel contrast exist in Chinese. While Korean has no long or short vowel contrast, Japanese and Vietnamese do have phonemic distinction between long and short vowels but Japanese still has no obstruent in its syllable codas (Saunders, 1987)

Significance of the current research

The current research aims to examine the acoustics correlates to vowel duration in two different contexts: in monosyllabic environment vs in word-isolation disyllabic environment where the target vowel lies on unstressed syllable and both environment end with voice contrasting obstruents.

The teaching of English as a second language in China, Korea, Japan, and Vietnam has mainly been focusing on an articulatory phonetics approach based on the contrastive analysis of first language and target language (Ref?). However, whilst the teaching of pronunciation is more focused, little attention is being devoted neither to provide students with the appropriate frame of reference for observing, imitating, and understanding the acquisition of pronunciation nor to test, to give feedback, or to fix their mistakes (Ref?).

In this study, both the effect of voicing contrasts and lexical stress on vowel duration were examined as it is once claimed that the stressed syllable is realized with increased F0 and longer duration than unstressed syllable (Fry, 1958; Gordon & Roettger, 2017).

The speech learning model (SLM)

It is stated that there are difficulties for adult learners in producing L2 contrasts and these difficulties may result from perceptual assimilation of both L1 and L2 sound systems. The production accuracy of L2 sounds might be limited to closest L1 properties. In support of this, Flege (1995) proposes the Speech Learning Model (SLM) to suggest people's sound system is reorganized when new sound categories are added. Best (1995) suggests Perceptual Assimilation Model (PAM) claiming that non-native speakers rely on their native phonemic systems when dealing with L2, which means if an L2 sound is similar to that in their L1 system, the sound will successively adopted. However, if an L2 sound is difficult to assimilate to the L1 sound category, they will have trouble acquiring that sound.

The following subsections review the literature on Chinese, Korean, Japanese, and Vietnamese pronunciation of English vowels in order to (1) investigate how different or

similar these languages share when learning English, (2) to compare and contrast some of the different descriptions and characteristics, from which procedures for future research are suggested that can help build a more practical description for those L2 speakers pronunciation of English in the future.

The English syllable

In English, the syllable structure is as below:
(C)(C)(C)V(C)(C)(C)(C)
The bare nuclear single vowel "V" can be the smallest possible syllable while up to three consonants can be included in the syllable onset and up to four syllables in its coda position. In the coda position, the consonant can be either voiced or voiceless (Giegerich, 1992).

The Chinese syllable

Chinese syllable follows this pattern:
(C)V(V)(V or N)
In Mandarin Chinese, one syllable corresponds to one morpheme. Each syllable is composed of three components: an onset, and a bimoraic rhyme. All onsets are single segments and the largest possible syllable can be CVVV or CVVN. When a syllable consists of a single vowel, the vowel is lengthened to fill the bimoraic rhyme. In Chinese, each syllable begins with a consonant and ends with a vowel, this can account for the situation that Chinese speakers of English have a tendency to add vowel after the consonant at coda position. For example, Chinese learners tend to produce the word "good" as / ˈgʊdə/, and "tap" as /ˈtæpʊ/. Unlike English, Mandarin Chinese has no voicing contrast for stops and stops only occur at the onset position. Chinese speakers of English have been shown to devoice voiced stops, delete voiced and voiceless stops, and also insert vowels after syllable final stops (Weinberger, 1987).

Similar to Japanese, Chinese is a syllable-timed language, which means the total number of syllable will determine the time to finish a Chinese sentence. As a result, Chinese speakers of English would produce equal length for each syllable when they speak English.

Relating to tone, Chinese is a tone language. Four tones in Mandarin Chinese are: high-level, high-rising, low-rising, and high-falling tone. It means a syllable can be pronounced in four different tones which express different meanings.

The Korean syllable

In Korean, the syllable structure is as below:
(C)V(C)
In Korean, there is no voiced consonants. All stops, fricatives, and affricates are all voiceless. Stress and vowel duration are not considered as lexically distinct. In Korean production of English, stress patterns and vowel duration are additional but not crucial to the meaning of lexical items. Korean is syllable-timed language (Lee and Jang, 2004; Kang, 2004; Yun, 2002) and the maximal syllable structure in Korean is CVC which means only one consonant is allowed in the onset and coda positions. Moreover, fricatives are not allowed in coda position, and consonants in coda position are never released in Korean (Berg and Koops, 2010).

The Japanese syllable

In Japanese, syllable structure follows this pattern:
(C)V(V or N)
Japanese syllables contrast due to the syllable weight (i.e light and heavy syllables) in which the unit of weight is called 'mora'. Light syllables have one mora and heavy ones have two morae. In Japanese, syllables consist of an onset, a nucleus and a coda, and like in English only the nucleus (i.e vowel) is obligatory (Iwasaki, 2002).

Relating to consonants, some syllables have onset while others do not. There are two types of codas in Japanese: coda closed with a nasal and those closed with a geminate consonant (i.e the consonant closes one syllable and serves as onset of the next syllable). As a geminate cannot occur at the word-final position, nasal is the only consonant that can serve as a genuine coda.

Syllable structure in Japanese is described as having no consonant clusters, nor obstruents in coda positions. To that end, isolated segments may be focused over segments in clusters, and this may affect the vowel intelligibility of Japanese people of English. Indeed, Saunders (1987) stated that vowel reduction was the favored process in Japanese Pronunciation of English (JPE).

The Vietnamese syllable

In Vietnamese, syllable structure follows this pattern:
(C)V(C)
Vietnamese is a monosyllabic language which means Vietnamese words are single-syllable words. Even though a word can have one or more syllable, each syllable is written separately. Even for loan words containing many syllables in their source language, the adaptation to monosyllabic form when written in Vietnamese must be followed. Vietnamese is one of the syllable-timed languages (Nguyen, 1970; Nguyen, 1980).

Maximal Vietnamese syllable structure is CVC. There is no consonant cluster at the beginning or at the end of a syllable. It means that in Vietnamese, no word can have the structure as CCVCC or CCCVCCC as in English. Moreover, the consonants on coda positions are not released (Nguyen et al, 2006).

However, Vietnamese syllable structure not only contains of consonants and vowels, but also tones. Similar to Chinese, tones in Vietnamese are used to change the meaning of the words. The syllable structure of Vietnamese was characterized by this table (Taiwan Buffalo International, 2001):

<table>
<tr><td rowspan="3">Initial</td><td colspan="3">Tone</td></tr>
<tr><td colspan="3">Final (rhyme)</td></tr>
<tr><td>Onset</td><td>Nucleus</td><td>Coda</td></tr>
</table>

Or more specific in Ngo (2005, p.7):

Tone			
Initial consonants	Labialization	Nuclear vowel	Final consonant/Semi-vowel

There are six consonants and two semi-vowels can be at word-final position. Those six consonants are nasal consonants /m, n, N/ and unaspirated voiceless plosives /p, t, k/. The two semi-vowels are /w/ and /j/ (Ngo, 2005).

Vowel lengthening before voiced coda in English

In English, one vowel phonemic length has two phonetically allophonic variations: short and long. It once has been noted that English vowels produced by an English native speaker lasted about 174 ms when they came before a voiceless consonant and 253 ms when preceding a voiced consonant (House and Fairbanks, 1953). A similar magnitude of 61ms difference of vowel duration in pre-voiceless versus pre-voiced position was claimed by Chen (1970).

House and Fairbanks (1955) described the environments in which the English vowel is lengthened, the table will be shown here.

Consonant Environment	**Average vowel length in milliseconds**
Fricatives	239
Nasals	232
Stops	203
Labio-dental consonants	234
Post-dental consonants	232
Bilabial consonants	205
Velar consonants	198
Voiced consonants	253
Voiceless consonants	174

Vowel lengthening: universal or language specific

As mentioned above, English has shown its pattern in the tendency that vowel is longer before voiced versus voiceless consonants. However, not only English but there are other languages that have been noted to have vowel duration distinctions before voiced/voiceless environment.

This table below will show the native vowel duration ratio across six languages, according to Chen (1970).

Language	**L1 Vowel Duration Ratio (VDR)**
English	.61
Korean	.78
Russian	.82
Spanish	.82
French	.87
German	.90

Among these languages, English has shown the lowest ratio of pre-voiceless to pre-voiced vowel duration, which means there has been an exaggeration of the tendency to lengthen the vowel before voiced coda than voiceless coda in English than other languages.

The data from above table might suggest that vowel lengthening before voiced consonant is a universal tendency. However, there have been studies relating to other languages such as Arabic, Polish, and Czech showed no significant difference of vowel length before a voiced coda.

Chinese vowels

Mandarin Chinese, which can be called Standard Chinese, has 5 vowels: /i/, /y/, /u/, /ə/, and /a/. The manner of articulation of Chinese and English vowels are different. Vowels are not lengthened before voiced codas, because vowels are of only one phonemic length in Chinese. For example, long vowel /ɪː/ and short vowel /ɪ/ in English are minimal pairs, however, there is no distinction of vowel /ɪ/ in Chinese. Because Chinese has no short /ɪ/ and long /ɪː/, speakers from Chinese L1 background may maintain the same length when articulating these sounds in different words.

Korean vowels

Korean vowels system contains two unrounded front vowels, which are /ɪ/ in high and /ɛ/ in non-high positions, non-fronted low vowel /ɑ/, two rounded vowels /u/ and /o/ at high back and mid back positions respectively. Lastly, there are high back unrounded vowel /ɯ/ and mid back unrounded vowel /ɤ/. Moreover, Korean vowel system also consists of ten diphthongs which are the combinations of glides and a vowel.

Japanese vowels

Japanese has been described as having five vowels /a, i, u, e, o/. All of them have functioned as separate phonemes under long forms. In a pitch-accented language as Japanese, the syllable on which the accent falls is considered more prominent than other syllables, which causes word-level distinction. In Japanese, the tone-bearing unit carries either high (H) or low (L) tones. However, not all words have accent, words having no accented syllable are called unaccented words. It has been stated that while English speakers rely predominantly on spectral properties for identifying English vowels (Hillenbrand, Clark, & Houde, 2000), Korean and Japanese speakers rely on the duration to discriminate vowel contrasts (Shibatani, 1990; Yang, 1996).

Vietnamese vowels

There are eleven vowels including long and short ones but not including diphthongs and triphthongs in Vietnamese. The duration of vowels can be considered as the primary cue for distinguishing minimal pairs (i.e /a/-/ɐ/, /ɤ/-/ʌ/). Vietnamese phonology describes ơ /ɤ/ and â /ʌ/ as long-short pair of a vowel, and similarly a /a/ and ă /ɐ/ forms a minimal pair.

Vowels and reinterpretation of distinction

To English native speakers, the phonemic difference between "bed" and "bet" is the voicing of final consonant, the voicing then can be classified as a distinctive signal since this is one of the ways to differentiate minimal pairs. As mentioned above, the length of vowel preceding final consonant can be another way to distinguish minimal pairs, and this can be considered as a redundant distinction.

When learning L2, learners try to acquire the ability to perceive and produce differences of L2 phonemes and allophones by transferring patterns from their L1. Guided by this principle, Vietnamese native speakers would be expected to use the same vowel before voiceless stop codas and devoice voiced stop codas. The limited codas as well as non-cluster language system have created difficulties for the Vietnamese in using voicing as a cue for distinguish minimal pairs.

The current study: Research questions and Hypothesis

Previous studies have shown data of vowel duration among speakers from different backgrounds other than English. However, the length of vowels of speakers from China, Korea, Vietnam, and Japan were not described. Do these systems in Asia share similar vowel lengthening patterns before voiced and voiceless codas? Are these patterns determined by characteristics of L1 vowel system, or by universal factors? Specifically, are these patterns language-specific or universal?

The purpose of this study is to describe the vowel duration produced by four groups of non-native English speakers, in order to examine the acquisition of the English allophonic vowel duration contrast by speakers of these four languages.

The study sought to answer two research questions:

- What are the English vowel length patterns before voiced and voiceless consonants among four subjects?
- Does the existence of a phonemic vowel length contrast in L1 language system help speakers to produce vowel length distinction in L2?

Hypotheses:

(1) Voicing effects are observed for monophthongs

(2) Stress and VQD in L1 affects L2 voicing effect more strongly than tone.

The participants in this study included speakers from four groups: native Chinese speakers (CN), native Korean speakers (KR), native Japanese speakers (JP), and native Vietnamese speakers (VN), and a baseline group of native English speakers (NS) to compare results from the four other groups. Non-native speakers from these groups came from different sociolinguistic backgrounds and there were many variables that need to be controlled when assessing their performance, however, despite the lack of sociolinguistic uniformity across these four languages, the interaction between the phonemic vowel contrast and speaker's ability to acquire L2 allophonic vowel length could still be examined.

Production experiment: Methodology

The production experiment aimed at measuring the English vowel durations of Chinese, Korean, Japanese, Vietnamese, and native English speakers, with the purpose of obtaining an average vowel duration ratio of each speaker group.

Participants

The participants in this study included 5 groups of adult speakers: native Chinese speakers, native Korean speakers, native Japanese speakers, native Vietnamese speakers, and native English speakers. Participants were recruited based on their native languages: Japanese (JP), Chinese (CN), Korean (KR), Vietnamese (VN), and US (NS). Five groups participated in the experiment, sectioned along the nationality. The group comprised 8 people from four different countries and 1 person from the US. They were all graduate students. Except for the NS, they are all non-native English speakers started their official school-based English education at around their teens in their home countries (China, Korea, Japan, and Vietnam). To that end, they all had more than ten years studying English. All subjects were recruited and paid for their participation. Participants were recruited by means of personal contact with the researcher by having responded to the letter sent via email of invitation to participate in a research.

The JP, CN, KR, and VN groups consisted of 8 speakers whose age ranges from 18-33, with origins from Tokyo, Beijing, Seoul, and Da Nang, respectively. The CN and KR reported at least some knowledge of a foreign language other than English. Beside English, they were fluent in Japanese. All speakers have normal speech and hearing functions with no history of communication disorders or intellectual impairments.

Procedures

The participants were instructed to pronounce each target word embedded in a carrier sentence. The speakers were provided with efficient time and instruction to complete the recording task. Each target word was presented 3 times. All the recordings were stored as 44kHz, 16-bit WAV files for manipulation.

Materials

Wordlist Stimulus

The stimuli used in this experiment were manipulations of real words. The choice was based on the fact that feedback from a lexical level to phonetic processing would not affect the result as the subject all had more than ten years of learning English, there will probably no significant difference in proficiency across groups.

The stimuli consist of common and rare words (APPENDIX B) in which common words can activate subject's lexical processing so that they will be produced with greater facility than rare words. Twenty-eight words under the form of minimal pairs, twelve disyllabic words in which target vowels are on unstressed syllable were recorded on a random basis for each subject by SONY PCM-D50 recorder and SONY ECM-959A microphone in a soundproof room. The total of forty words were put in a carrier sentence in order that no rising of falling tone might affect participants' production as it might be in single word list. Moreover, the

sentences containing target words were shown in a random order via PowerPoint slides so that the participants could not guess what would be going on. English was used as a language of instructions to activate the target language with each session lasted about ten minutes.

The distribution of vowels covers a range of vowel heights: one low vowel /ɒ/, one mid-vowel /æ/, one mid vowel /ɛ/, and one mid high vowel /ɪ/. Participants were asked to use the frame sentence "Say_____please" because the consonant-initial word following the target word was an environment that could prevent the linking of words.

Data analysis

For this experiment, the vowel duration of the 342 tokens were measured. These measurements were made using Praat software (Boersma and Weenink, 2017). Vowel duration was measured from the onset of periodicity showing clear formant and pulse to the end of periodicity signaled by a drop of amplitude in which the ending time referred to the zero-crossing point where the last pitch pulse ended followed by the final consonant; and the zero-cross point nearest to the vowel-to-consonant transition defined as the ending time of the vowel were all marked from spectrogram and waveforms using Praat 6.0.36 speech analysis software. Then, a Praat script was run to calculate vowel duration (Lennes, 2011).

Results

Figure 1 shows average vowel duration by language group. It can be seen that the, CN, VN, KR, and JP produced similar vowel duration before voiceless consonants compared to NS group. However, the vowel duration difference before voiced consonants across groups were fairly distinctive.

A statistical test showed a significant difference between voicing and vowel duration among tonal group, which is Chinese and Vietnamese ($p < 0.001$).

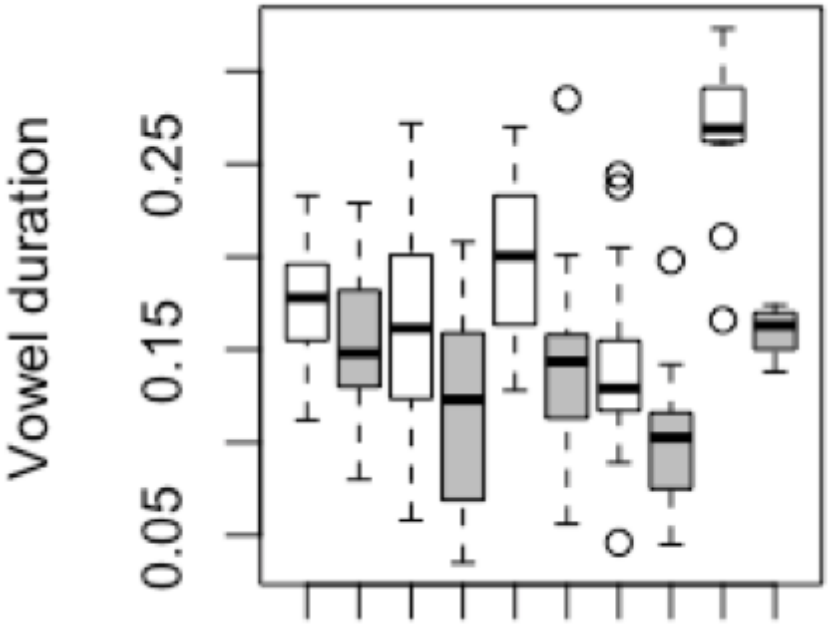

Fig. 1. Vowel duration followed by voiced/voiceless consonants, among 5 groups, monosyllabic (From left to right: CN, VN, KR, JP, and NS)

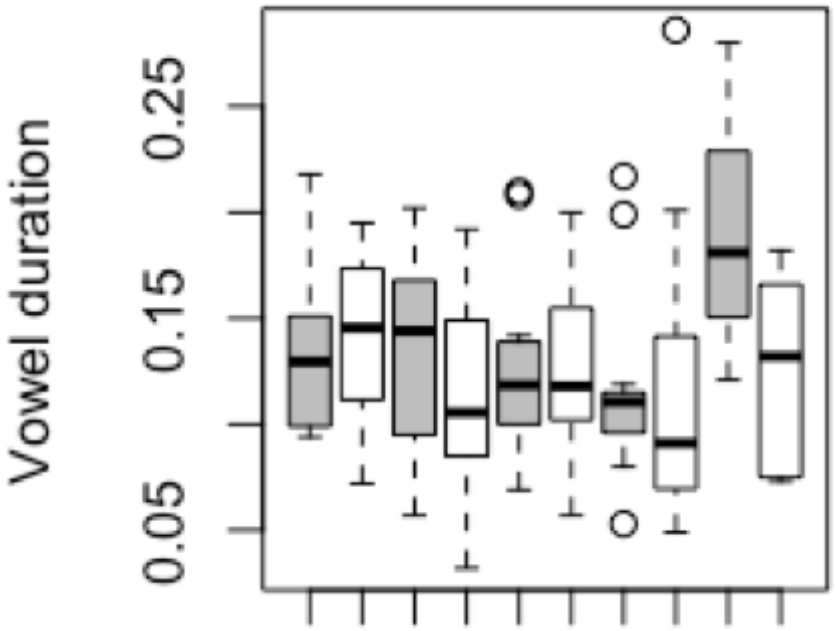

Fig. 2. Vowel duration followed by voiced/voiceless consonants, among 5 groups, disyllabic (From left to right: CN, VN, KR, JP, and NS)

The result of the data analysis procedure for the production experiment showed that the native English group a VDR between the range reported in previous studies which is 0.61 to 0.69. As can be seen from Fig.1 above and statistical data, the average vowel duration for each language differed slightly, according to whether it preceded voiced or voiceless consonant.

The first tonal group including Chinese and Vietnamese produced very similar vowel lengthening patterns in monosyllabic environment ($F= 0.95$, $p > 0.5$). It can be seen that all four groups showed similar pattern for longer vowel before voiced than voiceless codas. However, there was significant difference of vowel duration ratio of non-native compared to those of the native group ($F= 0.58$, $p < 0.01$).

However, in disyllabic environment, within tonal group, different vowel lengthening patterns of Chinese and Vietnamese speakers have been shown in which Chinese speakers lengthened vowel more before voiced than voiceless codas. This finding shows a lack of support for the hypothesis, indicating that Chinese and Vietnamese speakers may share similar patterns due to the tones those languages possess. More surprisingly, Korean group showed the same ratio for pre-voiced and pre-voiceless vowel duration.

Discussion and Conclusion

Results of monosyllabic words indicate our hypothesis 1 (universality) is true. However, the effect of tone, stress, and VQD are not consistently observed in our data, to that end, hypothesis 2 is not confirmed. The results of the data analysis procedure for the production experiment showed that disyllabic words showed inconsistent patterns which requires more extensive research. The results of the current study do raise some interesting questions.

There is one main factor that limited the generalizability of the results of the study: it was the small sample size which contributed to the assumption of data analysis in the experiment not being met. Future researchers should consider designing future experiments with a much larger sample size.

Appendices

APPENDIX A:SOCIOLINGUISTIC QUESTIONNAIRE

1. Where are you from?
2. What is your first language?
3. How long have you been learning English? Less than 1 year 1-2 years 2-5 years More than 5 years
4. What is your current age?
5. Do you speak any other languages. If so, list them below with your fluency level from 1 to 5 (1= beginner, 5= very fluent).

APPENDIX B: WORDLIST

Built-Build	Light-Lied	Cup-Cub
Lock-Log	Tote-Told	Bet-Bed
Sight-Side	Cap-Cab	Bailiff-Massive
Phallus-Pizzas	Racquet-Candid	Robot-Lingcod
Fallout-Newfound	Rowboat-Household	

References:

Best, C. T. (1995). A direct realist perspective on cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues in cross language speech research* (pp. 167–200). Timonium, MD: York Press.

Boersma, Paul & Weenink, David (2017). Praat: doing phonetics by computer [Computer program]. Version 6.0.36, retrieved 19 July 2018 from http://www.praat.org/

Chen, M. (1970). Vowel length as a function of the voicing of the consonant environment. *Phonetica: International Journal of Speech Science, 22,* 129-159.

Elsendoorn, B. A. G. (1983). Quality and quantity in English by Dutchman: Two parameters inducing double Dutch. In M. Van den Broecke, V. van Heuven, & W. Zonneveld (Eds.), *Sound structures, studiesfor Atonic Cohen* (pp. 53-69) Dordrecht, The Netherlands: Foris.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues* (pp. 233-277). Timmonium, MD.: York Press.

Flege, J. E., & Port, R. (1981). Cross-language phonetic interference: Arabic to English. *Language and Speech, 24(f),* 125-146

Fry, D. (1955). Duration and Intensity as Physical Correlates of Linguistic Stress. *The Journal Of The Acoustical Society Of America*, *27*(4), 765-768. doi: 10.1121/1.1908022

Giegerich, H. J. (1992). *English phonology: An introduction.* Cambridge, England: Cambridge University Press.

Gordon, M., & Roettger, T. (2017). Acoustic correlates of word stress: A cross-linguistic survey. *Linguistics Vanguard*, *3*(1). doi: 10.1515/lingvan-2017-0007

Hillenbrand, J., Clark, M., & Houde, R. (2000). Some effects of duration on vowel recognition. *The Journal Of The Acoustical Society Of America, 108*(6), 3013-3022. doi: 10.1121/1.1323463 House, A. S., & Fairbanks, G. (1953). The influence of consonant environment upon the secondary acoustical characteristics of vowels. *Journal o f the Acoustical Society ofAmerica,* 25(1), 105-114.

Kang, Y. 2004. *Understanding English Stress*. Seoul: Kyungjin.

Lee, J-P. & Jang, T-Y. 2004. A comparative study on the production of inter-stress intervals of English speech by English native speakers and Korean speakers. *ICSLP Proceedings*.

Mack, M. (1982). Voicing-dependent vowel duration in English and French: Monolingual and bilingual production. *Journal of the Acoustical Society of America, 71,* 173-178.

Ngo, Nhu Binh (2005) The Vietnamese learning framework – Part one: Linguistics. *COTSEAL Journal 2005.*

Nguyen, D. H. (1980). Language in Vietnamese society. Chicago: University of Illinois Press.

Nguyen, D. L. (1970). A contrastive phonological analysis of English and Vietnamese (Pacific Linguistics Series, No 8). Canberra: Australian National University.

Nguyen, Thien Giap and Doan, Thien Thuat et al. (2006). *Dẫn Luận Ngôn Ngữ Học.* [In Vietnamese]. Hà Nam: NXB Giáo Dục.

Saunders, N. (1987). Morphophonemic variation in clusters in Japanese English. Language Learning. 37.247-272.

Shibatani, M. (1990). The languages of Japan. Cambridge University Press, New York.

Taiwan Buffalo International (2001) *Sound system in Vietnamese*. http://www.de-han.org/vietnam/chuliau/lunsoat/sound/2.htm Weinberger, S. H. (1987). The influence of linguistic context on syllable structure simplification. In G. Ioup & S. H. Weinberger (Eds.), *Interlanguage phonology: The acquisition of a second language sound system* (pp. 401–417). Rowley, MA: Newbury House.

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

Yun, I. 2002. Compression effects of number of syllables on Korean vowel. *Speech Sciences*, 9(1): 173-184. The Korean Association of Speech Sciences.

Contact e-mail: n-le-2s3@eagle.sophia.ac.jp