INTERSPEECH 2017
August 20–24, 2017, Stockholm, Sweden

Effects of Pitch Fall and L1 on Vowel Length Identification in L2 Japanese

Izumi Takiguchi

Bunkyo Gakuin University, Japan

itakiguchi@bgu.ac.jp

Abstract

This study investigated whether and how the role of pitch fall in the first language (L1) interacts with its use as a cue for Japanese phonological vowel length in the second language (L2). Native listeners of Japanese (NJ) and L2 learners of Japanese with L1 backgrounds in Mandarin Chinese (NC), Seoul Korean (NK), American English (NE), and French (NFr) participated in a perception experiment. The results showed that the proportion of "long" responses increased as a function of vowel duration for all groups, giving s-shaped curves. Meanwhile, the presence or absence of a pitch fall within a syllable affected only NJ and NC's perception. Their category boundary occurred at a shorter duration for vowels with a pitch fall than without a pitch fall. Among the four groups of L2 learners, only NC use pitch fall to distinguish words in the L1. Thus, it is possible to think that the role of pitch fall as an L1 cue relates to its use as a cue for L2 length identification. L2 learners tend to attend to an important phonetic feature as a cue for perceiving an L1 category differentiating L1 words even in the L2 as implied by the Feature Hypothesis.

Index Terms: L2 speech perception, Japanese phonological vowel length contrast, L1 influence, perceptual cues

1. Introduction

Non-native contrasts in the L2 tend to cause problems to L2 learners. People agree that L1 relates to L2 learners' perception and/or perceptual difficulties even if it might not explain all of them. However, it remains unclear what property in the L1 affects L2 speech perception and what does not ([1]). The Feature Hypothesis ([1]) claims that phonetic features which are not exploited and are less prominent as cues for the L1 phonological contrasts are difficult to discern and to use in the L2. This seems to imply that phonetic features which are prominent in the L1 would accordingly be easy to use in the L2. In fact, L2 listeners are reported to use cues differently than native listeners and to rely on important L1 cues even in their L2 ([2], [3]).

Japanese has phonological vowel length contrast. Word meanings can be differentiated by vowel length alone, such as /kado/ (corner) with a phonemic short vowel versus /kado:/ (Japanese flower arrangement) with a phonemic long vowel. Vowel duration is the primary cue for the contrasting length categories ([4]). In addition to the length contrast, Japanese is known as a lexical pitch accent language, where the presence or absence of an accent on a particular syllable within a word differentiates words. An interaction between vowel length or duration and pitch has been reported. It is said that presence of a pitch fall within a syllable affects NJ's perception of vowel length and they tend to give more "long" responses to vowels with a pitch fall than without a pitch fall ([5], [6], [7], [8]).

Many L2 learners of Japanese tend to have difficulty in identifying phonological vowel length ([9]). [10] examined effects of pitch-related cues on identification accuracy of phonological vowel length using naturally spoken stimuli. The study found that Mandarin Chinese listeners identified Japanese long vowels with a pitch fall, i.e., HL as accurately as those with a high pitch, i.e., HH even though vowel duration of the stimuli was acoustically shorter for HL than for HH. Chinese listeners' results imply that they were not only listening to absolute vowel duration. If this were the case, length identification should be easier for HH than for HL due to the longer duration, but this was not the case. Therefore, it is possible to think that Chinese listeners heard the pitch fall on vowels and the pitch fall affected their perception of vowel length. Moreover, since Chinese is the language that uses pitch fall as a cue to differentiate words, the results indicate the possibility that the role of pitch fall in the L1 interacts with its use in the L2. However, few studies seem to have examined relation between the role of pitch fall for L1 phonological contrasts and its use in the L2 for length identification. Consequently, it is important to test the hypothesis that pitch fall affects vowel length identification and category boundary location only for those L2 listeners whose L1 employs pitch fall as a cue for L1 phonological contrasts.

The current study examined whether and how pitch fall in the L1 interacts with its use as a cue for vowel length contrast in L2 Japanese. A comprehensive approach for such an investigation needs to include collecting and comparing data from those L2 learners of Japanese whose L1 differs in terms of pitch fall as a cue. Accordingly, based on the role of pitch fall as an L1 cue, native listeners of Mandarin Chinese, Seoul Korean, American English, and French were selected. Mandarin Chinese uses pitch fall for lexical tones. American English uses pitch fall for lexical stress. Seoul Korean does not exploit pitch fall to distinguish one word from another even though it is reported to use pitch height for three-way laryngeal contrasts. French, in contrast, does not exploit pitch related cues to differentiate words. This study explored whether and how these differences in the L1 affect listeners' use of pitch fall for identification of vowel length in the L2.

2. Methodology

2.1. Participants

There were 56 participants in total: nine NJ (control group; 2 males and 7 females), 11 NC (5 males and 6 females), 12 NK (3 males and 9 females), 12 NE (5 males and 7 females), and 12 NFr (8 males and 4 females). The nonnative listeners of Japanese were all adult L2 learners, who started learning Japanese after the age of 15 years and had previously taken or were taking formal instruction in Japanese.

The NJ group's ages ranged from 19 to 48 years with the mean age being 29.9 years (SD = 11.2). All of the participants

http://dx.doi.org/10.21437/Interspeech.2017-763

in the NC group are from mainland China and reported that they spoke Mandarin Chinese as their L1. Their ages were between 20 and 28 years with the mean age being 23.5 years (SD＝2.4). The participants in the NK group were all born and grew up in Seoul. Their ages ranged from 20 to 32 years with the mean age being 22.9 years (SD＝3.4). The NE were all born and grew up in the US. Their mean age was 20.7 years (range: 18-25, SD＝1.7). Finally, the NFr were all born and grew up in France. Their mean age was 23.3 years (range: 19-31, SD = 3.6). No participants reported any hearing or speaking disorder.

2.2. Stimuli

All stimuli were created from a token of a nonsense word /nono:/ produced with an HHH accent pattern, which is an impossible accent pattern in Tokyo Japanese, in isolation. A female native speaker of Tokyo Japanese in her twenties pronounced /nono:/ in an HHH accent pattern 20 times and in LHL and LHH patterns 12 times in a random order. The tokens were recorded using a linear PCM recorder (SONY PCM-D50) at a 44.1 kHz sampling rate and 16-bit quantization with an electric condenser microphone (SONY ECM-MS957). From the 20 tokens with an HHH pattern, one token which had duration similar to the mean was selected for stimulus manipulation for the experiment. Table 1 summarizes acoustic characteristics of the selected token.

Table 1: *Segment durations (ms) and F0 values (Hz) of the original token.*

segment		n	o	n	o:
duration		26.4	96.3	70.6	339.6
F0	onset	235.6	226.4	226.4	227.5
	offset	226.4	226.4	227.5	209.5

F0 and timing information of tonal events were extracted from the 12 naturally spoken tokens for the two accent patterns, LHL and LHH to use them as reference data for manipulation. Based on the J_ToBI convention introduced as a prosodic labeling system for the description of Japanese prosodic structure by [11], F0 values of boundary tones, F0 values at the onset of each segment along with the max F0 value of the utterance and its locus were measured for each token using Praat ([12]). Table 2 shows mean F0 values and mean time of the locus of max F0.

Table 2: *Mean F0 values (Hz) and mean time (second (s)) of the location of max F0 in a word.*

	%L	C1 onset	V1 onset	C2 onset	V2 onset	L%	Max F0 & the loci
LHL	156.9	214.9	191.1	188.0	230.7	156.1	249.9 0.25866
LHH	175.2	206.8	182.1	180.8	210.9	209.7	220.5 0.29612

A total of 30 stimuli were created by manipulating V2 duration and the accent pattern of the original token using the PSOLA function in Praat. First, noise before C1 and after V2 was cut off and 20 ms-silence was inserted instead. Next, the accent pattern was manipulated by removing all pitch points of the original token and setting new pitch points as the F0 values as shown in Table 3. After manipulating pitch contour, V2 duration for each accent pattern was shortened or lengthened from 40 ms to 390 ms in 25 ms steps (15 steps). The durations of other parts except V2 duration were kept the same as in the original token. This manipulation yielded 30 stimuli (15 steps × 2 accent patterns), which were identical in the durational distribution, but varied in the accent pattern, namely LHL or LHH. The stimuli were checked by two trained phoneticians and were judged as natural Japanese sounds /nono/ or /nono:/. Also, the first syllable was always heard as having a short vowel.

Table 3: *Manipulation of the LHL and LHH patterns: Loci of pitch points in time (s) and F0 values (Hz).*

	C1 onset	V1 onset	C2 onset	V2 onset	max	max	L%
Pitch points	0.019962	0.046315	0.142616	0.213236	0.278619	0.316174	0.552823
LHL	214.9	191.1	188.0	230.7	249.9	-	156.1
LHH	206.8	182.1	180.8	210.9	-	220.5	209.7

2.3. Procedure

Each stimulus was presented to participants eight times in a random order, so the test section consisted of 240 trials (15 stimuli × 2 accent patterns: LHL and LHH × 8 times). The trials were divided into 12 blocks. The participants did a practice section containing eight trials (2 endpoints × 2 accent patterns × twice) prior to the test section. SuperLab [Cedrus, ver. 4.5] was used to present stimuli to the participants and to collect responses from them. Participants were asked to identify whether the word was /nono/ or /nono:/. They were told to press one of the two designated keys on the PC, which were labelled as “nono” and “nono:” in Japanese *katakana*, as quickly as possible within three seconds. The next stimulus was presented after participants responded or the 3,000 ms response period had elapsed, whichever came first. It took approximately 8-15 minutes for participants to complete the experiment. All participants were paid for their participation.

3. Results

Figure 1 shows the logistic functions for each L1 group for each accent pattern. The lines and the shapes represent the fitted logistic curves and the observed proportion of “long” responses, respectively: the broken lines and triangles for LHL and the solid lines and cross marks for LHH. The proportion of “long” responses increased as a function of V2 duration, giving an s-shaped curve for all groups for both accent patterns. The logistic curves demonstrated by NJ and NC are similar in that the two curves separate each other and the curve for LHL locates to the left of that for LHH. This indicates that both NJ and NC gave more “long” responses to vowels with a pitch fall than those without a pitch fall. NFr’s curves are also apart and the curve for LHL locates to the left of the curve for LHH but to a lesser degree compared to those of NJ and NC. In contrast, NK’s two logistic curves overlap, which implies that the presence or absence of a pitch fall did not affect their length identification. In addition, we see that NE’s logistic curves cross at around the 50 % cross-over point.

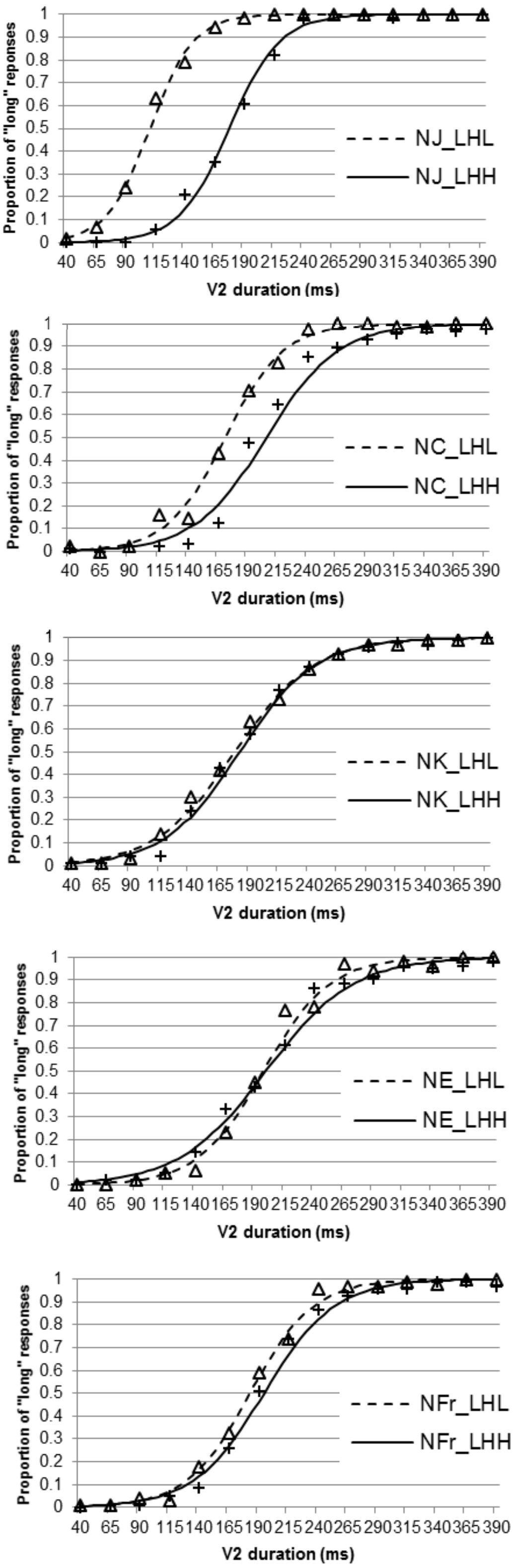

Figure 1: *Proportion of "long" responses as a function of vowel duration for LHL and LHH. From the top to the bottom: NJ, NC, NK, NE, and NFr.*

Figure 2 shows mean estimates of the category boundary, which was estimated as the 50% cross-over point of response rate on the fitted logistic functions, for each L1 group for each accent pattern, striped bars for LHL and white bars for LHH. As we could guess from the logistic curves in Figure 1, NJ and NC clearly demonstrate the shorter category boundary value for LHL than for LHH. NFr shows the same tendency as NJ and NC but to a lesser degree. The category boundary values of NE and NK do not seem to be different between the two accent patterns. A two-way repeated measures ANOVA for the data on the category boundary location with L1 (NJ vs. NC vs. NK vs. NE vs. NFr) as a between-subjects factor and accent pattern (LHL vs. LHH) as a within-subjects factor showed significant main effects of L1 [$F(4, 51) = 6.7$, $p<.001$] and accent pattern [$F(1, 51) = 39.8$, $p<.001$]. In addition, there was a significant interaction between L1 and accent pattern [$F(4, 51) = 9.4$, $p<.001$]. Since an interaction between L1 and accent pattern was significant, a one-way repeated measures ANOVA was conducted for each L1 group with accent pattern as a within-subjects factor. ANOVAs revealed that there was a significant main effect of accent pattern for NJ [$F(1, 8) = 48.3$, $p<.001$] and NC [$F(1, 10) = 11.1$, $p<.01$] but that there was no main effect of accent pattern for the other three groups, NK [$F(1, 11) = 0.31$, $p=0.592$, *n.s.*], NE [$F(1, 11) = 0.09$, $p=0.768$, *n.s.*], and NFr [$F(1, 11) = 2.9$, $p=0.114$, *n.s.*]. The results showed that the category boundary of LHL was significantly smaller than that of LHH for NJ (109.1 ms vs. 176.9 ms) and NC (170.5 ms vs. 205.3 ms). In contrast, the two boundaries did not significantly differ from each other for the other three groups, NK, NE, and NFr.

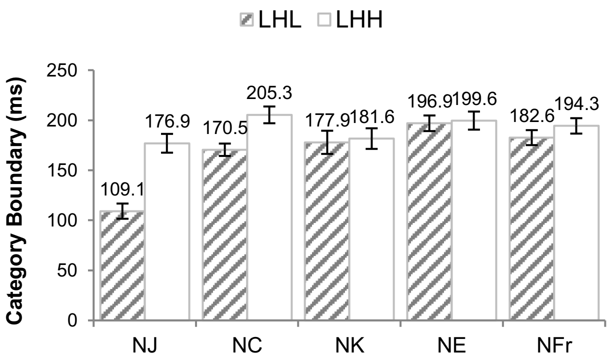

Figure 2: *Mean estimates of the category boundary.*

4. Discussion

It is possible to say L2 learners, irrespective of their non-use of vowel duration for phonological vowel length in the L1, can access vowel duration as an L2 cue. The proportion of "long" responses increased as a function of V2 duration in a rather similar fashion for all groups. These similar s-shaped curves can be interpreted as showing that all groups of listeners identify vowel length primarily based on vowel duration even though none of their L1s employs vowel duration as a cue for vowel length contrast. Non-use of vowel duration for L1 phonological vowel length does not result in a loss of ability to use vowel duration as a cue in the L2 ([13]); however, obviously, this does not necessarily mean that all L2 listeners regardless of the L1 can use vowel duration as NJ do ([14]).

Pitch fall serves as a secondary cue for NJ's length identification and the association between pitch fall and long vowels seems to relate to the shorter category boundary value for vowels with a pitch fall. Pitch fall within a syllable affected NJ's identification of vowel length, which is consistent with the results in previous studies ([5], [6], [7], [8]) in that Japanese listeners use pitch fall as a cue for vowel length. We can clearly see that NJ were biased toward giving more "long" responses by the presence of a pitch fall on a vowel. In fact, NJ's boundary location occurred at a significantly shorter duration for vowels with a pitch fall. [6] and [7] suggest that the bias toward the perception of long vowels on vowels with a pitch fall stems from the distributional fact that a pitch fall can occur only on long vowels in the language. That is, the presence of a pitch fall on a vowel associates with two moras, i.e., long vowels, for Japanese listeners and such association affects their perception.

Among L2 learners, only NC's length identification was affected by the presence of a pitch fall on a vowel. What experience in their L1 made NC use pitch fall and contrastively led the other three groups of listeners not to use it as a cue? We can think that NC attend to a phonetic feature that is important as a cue in their L1 even when listening to the L2, as implied by the Feature Hypothesis ([1]). For NC, the presence or absence of a pitch fall within a syllable distinguishes two tones, Tone 1, where a high pitch lasts from the beginning to the end, and Tone 4, where a high pitch at the beginning falls to a low pitch at the end. This suggests that the presence or absence of a pitch fall within a syllable associates with one specific category or another and in turn functions as an important cue to perceive the contrast. Such L1 experience leads NC to draw an attention to pitch fall and use it as a cue in the L2. NC transferred their experience with pitch fall, which is an important tonal difference in their L1 to discriminate one category from another that differentiates words, to L2 perception and relied on pitch fall as a cue, as implied by the Feature Hypothesis.

On the other hand, non-use of pitch fall for L2 length identification by NE, NK, and NFr can be explained by the insignificance of pitch fall as a cue for a specific category distinguishing one word from another in each of the three languages. Pitch movement seems to function as a cue for lexical stress in American English. According to previous studies ([15], [16]), English listeners are more likely to perceive the presence of lexical stress when pitch movement occurs on a vowel. However, it is said that pitch is exploited as a cue for lexical stress only when lexical stress coincides with sentence stress ([17], [18], [19], [20]). In addition, lexical stress in English can be realized with several pitch patterns. Thus, pitch fall is just one of them and does not necessarily associate with the presence of lexical stress. In short, the role of pitch fall may be weaker in American English than in Mandarin Chinese in that its presence or absence does not always associate with a specific category in the former as it does in the latter. In Seoul Korean, whether pitch is high or low is said to be an important cue for the three-way laryngeal contrast among stop consonants, especially between the lenis and aspirated series, where the primary cue, VOT significantly overlaps. However, there seems to be no report on the use of pitch fall as a cue in the language. As for French, pitch is used as a cue for the obligatory primary stress and the final full syllable of a word often bears a rising pitch movement when it is the last full syllable of a non-sentence final phrase ([21]). However, French is known as a fixed stress language and the primary stress always falls on the last full syllable of the final word in a phrase ([22], [23], [24]). This means that in French pitch does not differentiate any words, and pitch fall has no lexically distinctive function. We may argue that the difference in the role of pitch fall in the L1 is reflected in NC's use and NE, NK, and NFr's non-use of pitch fall in the L2.

NC's category boundary occurred at a significantly shorter duration for vowels with a pitch fall (170.5 ms) than for those without a pitch fall (205.3 ms), which is the same direction of the boundary shift as made by NJ. We can think of three possible explanations for this category boundary shift. First, it is possible that Mandarin Chinese has some characteristics which lead NC to perceive longer duration on vowels with a pitch fall than without a pitch fall. The stimuli with a LHL pattern would sound similar to Tone 4 (high falling tone) in that both have pitch fall within a syllable. Previous studies agree that Tone 4 is the shortest among the four lexical tones in Mandarin Chinese ([25], [26], [27]). Thus, if NC use average duration associated with lexical tones as a criterion for L2 vowel length identification, we can expect that their durational criterion for vowels with a pitch fall would be shorter than that for those without a pitch fall, which sounds similar to Tone 1 (high level tone). Then if they give a response as a long vowel when they think that the presented vowel duration is above the criterion, their category boundary for LHL should occur at a shorter duration than that for LHH, which was what we observed here. Second, it might be the case that NC gained phonological knowledge that pitch fall can occur on long vowels but not on short vowels through L2 learning of Japanese and that this led them to use pitch fall as a secondary cue for length identification. Since Mandarin Chinese is a lexical tone language, NC have experience with pitch fall as a cue for distinct lexical tones. This suggests that NC are able to hear variations in accent patterns in Japanese. Such an ability to hear pitch-related differences might allow NC to notice and to learn the distribution of possible pitch patterns on short and long vowels in Japanese. Thus, NC use vowel duration as the primary cue and pitch fall as a secondary cue for identification of vowel length, the same as NJ do. Finally, NC's longer perceived duration on vowels with a pitch fall might have stemmed from a universal-type of interaction between pitch movement and perceived duration in speech perception. Previous studies ([28], [29], [30], [31]) agree that perceived duration tends to be longer when stimuli carry pitch movement compared to when they do not and this is assumed to be present in speech perception universally. Thus, current NC's result might have reflected such general interaction between pitch movement and perceived duration. Further study should be conducted to verify the hypotheses.

5. Conclusions

The current study investigated how pitch fall in the L1 interacts with its use as a cue for vowel length in L2 Japanese. The importance of pitch fall as a cue differs between Mandarin Chinese and the other three languages. This difference in the role of pitch fall as a cue was reflected in listeners' use and non-use of pitch fall for L2 length identification. As Chinese listeners show, L2 listeners are likely to attend to a phonetic feature that is an important cue to perceive a category differentiating L1 words even when listening to the L2, which is implied by the Feature Hypothesis. To conclude, the role of phonetic features in the L1 determines which L2 acoustic information listeners pay attention to.

6. Acknowledgements

An earlier version of this paper appeared as part of my Ph.D. dissertation submitted to Sophia University in 2014.

7. References

[1] R. McAllister, J. E. Flege, and T. Piske, "The Influence of L1 on the Acquisition of Swedish Quantity by Native Speakers of Spanish, English and Estonian," *Journal of Phonetics*, vol. 30, no. 2, pp. 229-258, 2002.

[2] P. Lippus, K. Pajusalu and J. Allik, "The Tonal Component of Estonian Quantity in Native and Non-native Perception," *Journal of Phonetics*, vol. 37, no.4, pp. 388-396, 2009.

[3] Q. Wang, "L2 Stress Perception: The Reliance on Different Acoustic Cues," in *the 4th International Conference on Speech Prosody*, *Campinas, Brazil, proceedings*, Retrieved from http://www.isca-speech.org/archive/sp2008/papers/sp08_35.pdf, 2008.

[4] H. Fujisaki, K. Nakamura, and T. Imoto, "Auditory Perception of Duration of Speech and Non-speech Stimuli," in G. Fant and M. A. A Tatham, *Auditory Analysis and Perception Speech*. London: Academic, pp. 197-219, 1975.

[5] K. Kinoshita, D. M. Behne, and T. Arai, "Duration and F0 as Perceptual Cues to Japanese Vowel Quantity," in *the 7th International Conference on Spoken Language Processing*, *Denver, USA, Proceedings,* Retrieved from http://www.splab.net/papers/2002/2002_03.pdf, 2002.

[6] H. Lehnert-LeHouillier, "A cross-linguistic investigation of cues to vowel length perception," *Journal of Phonetics*, vol. 38, no. 3, pp. 472-482, 2010.

[7] Y. Nagano-Madsen, *Mora and Prosodic Coordination: A Phonetic Study of Japanese, Eskimo and Yoruba*. Lund, Sweden: Lund University Press, 1992.

[8] I. Takiguchi, H. Takeyasu, and M. Giriko, "Effects of a Synamic F0 on the Perceived Vowel Duration in Japanese," in *the 5th International Conference on Speech Prosody, May 11-14, Chicago, USA, Proceedings*, 100944, pp. 1-4, 2010.

[9] T. Toda, "Gaikokujin Gakushuusha no Nihongo Tokushuhaku no Shuutoku [Acquisition of Special Morae in Japanese as a Second Language]," *Journal of the Phonetic Society of Japan*, vol. 7, no. 2, pp. 70-83, 2003.

[10] I. Takiguchi, *Duration and Pitch Cues in L1: Perceptual Effect on Phonological Vowel Length for Learners of Japanese*, Doctoral dissertation, Sophia University, 2014.

[11] J. J. Venditti, "The J_ToBI Model of Japanese Intonation," in S. –A. Jun, *Prosodic Typology: The Phonology of Intonation and Phrasing*. New York, NY: Oxford University Press, pp. 172-200, 2005.

[12] P. Boersma and D. Weenink, "Praat: Doing Phonetics by Computer (Version 5.3.11) [Computer program]," Retrieved from http://www.praat.org/ (April 1, 2012), 1992-2012.

[13] I. Takiguchi, "Effects of Pitch Height on L2 Learners' Identification of Japanese Phonological Vowel Length," in S. Tanaka, G. Pinter, S. Ogawa, M. Giriko, and H. Takeyasu, *On'in Kenkyu no Shintenkai: Kubozono Haruo Kyooju Kanreki Kinen Ronbunshuu*. Kaitakusha: Tokyo, pp. 295-311, in press.

[14] I. Takiguchi, "The Role of Vowel Duration Cue in L1: Effects on L2 Learners' Identification of Phonological Vowel Length in Japanese," in *the 18th International Congress of Phonetic Sciences, August 9-14, Glasgow, UK, Proceedings,* Paper number 0111, pp. 1-5, 2015.

[15] Cutler, "Lexical Stress," in D. B. Pisoni and R. E. Remez, *The Handbook of Speech Perception*. Malden, MA: Blackwell Publishing, pp. 264-289, 2005.

[16] D. B. Fry, "Duration and Intensity as Physical Correlates of Linguistic Stress," *Journal of the Acoustical Society of America*, vol. 27, no. 4, pp. 765-768, 1955.

[17] M. Ortega-Llebaria, H. Gu, and J. Fan, "English Speakers' Perception of Spanish Lexical Stress: Context-driven L2 Stress Perception," *Journal of Phonetics*, vol. 41, no. 3-4, pp. 186-197, 2013.

[18] I. Shport, *Cross-linguistic Perception and Learning of Japanese Lexical Prosody by English Listeners*, Doctoral dissertation, University of Oregon, Retrieved from http://pages.uoregon.edu/ishport/Shport_Diss_2011_CrossLingLexicalProsody.pdf, 2011.

[19] A. M. C. Sluijter and V. J. van Heuven, "Acoustic Correlates of Linguistic Stress and Accent in Dutch and American English," in *the 4th International Conference on Spoken Language Processing, Philadelphia, USA, Proceedings*, pp. 630-633, Retrieved from http://www.asel.udel.edu/icslp/cdrom/vol2/604/a604.pdf, 1996.

[20] M. Ueyama, *Prosodic Transfer: An Acoustic Study of L2 English vs. L2 Japanese*, Doctoral dissertation, University of California at Los Angeles, Retrieved from http://www.linguistics.ucla.edu/facility/research/ ueyama.pdf, 2000.

[21] S. –A. Jun and C. Fougeron, "Realization of Accentual Phrase in French Intonation," *Probus*, vol. 14, pp. 147-172, Retrieved from http://www.linguistics.ucla.edu/people/jun/Probus-French.pdf, 2002.

[22] A. Di Cristo, "Intonation in French," in D. Hirst and A. Di Cristo, *Intonation Systems: A Survey of Twenty Languages*. Cambridge, UK: Cambridge University Press, pp. 195-218, 1998.

[23] A. M. Kijak, *How Stressful is L2 Stress?: A Cross-linguistic Study of L2 Perception and Production of Metrical Systems*, Doctoral dissertation, Utrecht University, Netherlands, Retrieved from http://igitur-archive.library.uu.nl/dissertations/2009-0615-200413/kijak.pdf, 2009.

[24] Tranel, *The Sounds of French: An Introduction*. New York, NY: Cambridge University Press, 1987.

[25] A. Jongman, W. Yue, C. B. Moore, and J. Sereno, "Perception and Production of Mandarin Chinese Tones," in P. Li, L. H. Tan, E. Bates and O. J. L. Tzeng, *Handbook of East Asian Psycholinguistics: Vol. 1, Chinese*. Cambridge, UK: Cambridge University Press, pp. 209-217, 2006

[26] Y. Xu, "Contextual Tonal Variations in Mandarin," *Journal of Phonetics*, vol. 25, no. 1, pp. 61-83, 1997.

[27] Y. Zhang, S. L. Nissen, and A. L. Francis, "Acoustic Characteristics of English Lexical Stress Produced by Native Mandarin Speakers," *Journal of the Acoustical Society of America*, vol. 123, no. 6, pp. 4498-4513, 2008.

[28] I. Lehiste, "Influence of Fundamental Frequency Pattern on the Perception of Duration," *Journal of Phonetics*, vol. 4, no. 2, pp. 113-117, 1976.

[29] D. B. Pisoni, "Fundamental Fequency and Perceived Vowel Duration," *Journal of the Acoustical Society of America*. Vol. 59, no. S39, 1976.

[30] W. S. –Y. Wang, I. Lehiste, C. –K. Chuang, and N. Darnovsky, "Perception of Vowel Duration," *Journal of the Acoustical Society of America*, vol. 60, no. S92, 1976.

[31] A. C. L. Yu, "Tonal Effects on Perceived Vowel Duration," in C. Fougeron, B. Kühnert, M. D'Imperio, and N. Vallé, *Laboratory Phonology 10*. Berlin: Mouton de Gruyter, pp. 151-168, 2010.