Perception of lexical pitch-accent by Korean Learners of Japanese

Shuangshuang Hu

UiL-OTS, Utrecht University
s.hu@students.uu.nl

Abstract

Tokyo Japanese has a lexical pitch-accent system whereas Seoul Korean features no word-level tonal representations. The present study investigated phonological perception of lexical pitch-accent by advanced Seoul Korean learners of Japanese, via sequence-recall experiments. The results, on the contrary to our prediction, showed that advanced Korean learners performed as well as native Japanese speakers in perceiving pitch-accent contrasts. Non-significant variance also suggested the advanced Korean learners in general have acquired lexical pitch-accent phonologically. The observations that Korean learners had no difficulties in perception of pitch-accent were discussed with respect to acquisition models of non-native perception.

Keywords: non-native speech perception, pitch accent, L2 acquisition, phonological perception

1. Introduction

Learners of foreign languages have difficulties in processing the target language in many aspects. Phonology is one of these. Previous studies have shown second language learners have encountered difficulties in parsing segmental (e.g., vowels and consonants) and phonological features (e.g., stress). For instance, difficulty in discriminating /r/ vs. /l/ is persistent in Japanese learners of English [1]. Advanced French learners of Spanish have 'persistent deafness' in perceiving lexical stress [2]. Theoretical models have been proposed to account for these perceptual difficulties that the lack of certain phonological representations in the native language influences non-native perception [3-4].

Empirical investigations on Seoul Korean (Standard Korean) prosody have suggested that the delimitation of prosodic units is mainly determined by the tonal patterns or pitch contours in the domain of accentual phrase – an intonational unit, a phrasal-level instead of word-level [5]. Different to Seoul Korean, the language with null phonological representations in lexical level, Tokyo Japanese has a lexical pitch-accent system. The meaning of a word is determined by where a pitch fall from high to low occurs in multi-syllable [6]. Words are presented in morae realized by low pitch (L) or high pitch (H) in an accentual pattern. For instance, disyllabic word /haʃi/ has three meanings depending on the location and presence of pitch-accent: 'chopsticks' in H'-L, 'bridge' in L-H' and 'edge' in L-H.

Given the prosodic properties of the two languages, previous L1 and L2 perceptual studies have focused on examining whether Korean speakers can perceive segments in phonetic repertoire or a particular pitch cue in Japanese prosody. Kim [7] found native Seoul Korean speakers were not sensitive to word initial voiced and voiceless consonants regardless of pitch variation in H'-L vs. L-H' patterns. Takiguchi [8] investigated perception of word-final vowel length by Korean speakers and found pitch duration did not affect their perception. Studies on the perception of pitch-accent by Korean learners of Japanese have centered on whether they were able to perceive the location of pitch-accent in different syllables and these studies applied same approach: accent listening test to mark where the pitch-accent locates [9-10].

Different from previous perceptual studies that applied acoustic approaches, the present study aimed to investigate phonological perception of lexical pitch-accent by Korean learners of Japanese. We applied a sequence-recall task developed by Dupoux et al. [2, 11]. Some models predict that the lack of some abstract phonological representations in the native language result in an impairment in non-native perception [2, 3-4]. Seoul Korean is not a stress language, nor a tone language or lexical pitch-accent language. Pitch-accent contrasts do not exist in Korean. Hence, we predicted the advanced Korean learners of Japanese had difficulties in perceiving lexical pitch-accent.

2. Method

2.1 Subjects

Twenty four subjects, 12 Japanese native speakers and 12 Seoul Korean learners of Japanese, participated in the experiment. They were divided into two groups, Japanese as control group and Korean as target group. Subjects in Japanese group (2 females and 10 males, aged from 21 to 30, average age of 25.8) came from the areas: Saitama, Kanagawa, Tokyo and Chiba, where speak Tokyo Japanese (standard Japanese) [12]. Subjects in Korean group consisted of undergraduate and graduate students from University of Tokyo, Tokyo University of Foreign Studies and The University of Electron-Communication. Among the Korean subjects, two were from Department of Japanese Language Studies in Tokyo University of Foreign Studies while the others were from non-linguistic/language studies. All of the Korean subjects (6 females and 6 males, aged from 23 to 32, average age of 26.1) came from areas in Gyeonggi Province including Seoul, where speak Seoul Korean. All of the Korean subjects have passed Japanese Language Proficiency Test N1 (advanced) and have stayed in Tokyo for an average of 3 years.

2.2 Stimuli

Bimoraic non word /nono/ in different accentual patterns was used. There are three possible accentual patterns of bimoraic words: H'-L (initial mora accented), L-H' (final mora accented) and L-H (unaccented), which differ in the position and the presence of where the pitch accent falls. L-H' and L-H are minimal pairs, which differ on the presence of pitch-accent on the final mora when a word-level grammatical word follows [13-14]. The three accentual patterns result in three contrasts: H'-L vs. L-H', H'L vs. L-H and L-H' vs. L-H as illustrated in Table 1. All the tokens of /nono/ were extracted from the sentence "kinoo /nono/ ga kita" (yesterday /nono/ came).

Table 1. Pitch-accent contrasts used in the experiment.

Contrast	First Item	Second Item
Initial accented vs. Final accented	H'-L	L-H'
Initial accented vs. Unaccented	H'-L	L-H
Final accented vs. Unaccented	L-H'	L-H

Several tokens of /nono/ were recorded by a trained Japanese female phonetician. Maximal and minimal f0 values were controlled, peak intensity of the H accented tone were set to the same value. Tokens of lengthened vowel and glottalized final mora were eliminated. The average duration of the tokens was 96.1 ms.

2.3 Procedure

The experiment was modified on the paradigm of Doupox et al. [11] and was conducted on a laptop with a button box, programmed in ZEP. There were three parts in the experiment, each part containing each contrast. The order of the parts was counterbalanced. Each part consisted of three phases: learning phase, practice phase and test phase. The subjects listened to the stimuli through headset in all the parts. Five minutes' break was between each part.

The learning phase consisted of two blocks, a learning block and a training block. Subjects were first instructed to learn new words in a foreign language in the learning block. They heard the first item (e.g., token /nono/ in H'-L) by pressing button A, and the second item (e.g., token /nono/ in L-H') by pressing button B. They were required to associate the buttons with the two stimuli. They could press the two buttons and listen to as many tokens as they wished until they were ready to proceed in a training block. They were asked to identify the stimuli they heard by pressing the corresponding buttons A or B. During our pilot tests, we found in the contrast of L-H' vs. L-H some subjects failed to achieve seven-time correct responses in a row, the criterion used in Dupoux et al. [11]. Thus the criterion in the training block was modified to a success of 12 correct responses among the total 18 trials. The training block would continue until the subject reached the criterion.

The following practice phase consisted of one block. During this phase, the subjects listened to a two-word (e.g., AB) and three-word (e.g., ABA) of the stimuli and were asked to press the buttons corresponding to the stimuli in the sequence. The sequences of the stimuli were presented in random, each sequence for four trials in total eight trials. Each trial was 1500ms, with an ISI of 80 ms included, and was separated by the word "ok". A feedback was shown on the screen for 700 ms after each trial.

In the test phase, two-, three- and four-word sequence were presented in three blocks. The

design was in general the same as in practice block, but without a feedback and a four-word sequence was included in the test phase. Each of the four possible sequences was used twice in two-word sequence block. In three-word sequence block, each of the eight possible sequences was used once. In four-word sequence block, eight possible sequences were chosen and each was used once. Thus the test phase contained 24 trials.

The experiments for the Korean group were carried out in a sound proof room or a quite room in University of Tokyo, Tokyo University of Foreign Studies and The University of Electro-Communications; for Japanese group were tested in National Institute for Japanese Language and Linguistics in Tokyo.

2.4 Data Analysis

Accuracy rate for each contrast and each block (two-, three- and four-word sequence) in each subject were analyzed. Mean accuracy rates of each block in each contrast were transformed to z-score. Based on Van Selst and Jolicoeur [15], the subject whose accuracy rate was beyond the range of (-2.246, 2.246) was regarded as an outlier. Data of one subject (a z-score: -3.18) in Korean group and one subject (a z-score: -3.20) in Japanese group were eliminated. Statistics analysis was conducted as stated below.

3. Results

Accuracy rates of all the subjects were analyzed via Repeated Measures in ANOVA with one between-subject factor (Korean vs. Japanese), and two within-subject factors: Contrast (H'L vs. LH', H'L vs. LH, LH' vs. LH) and sequence length (two vs. three vs. four). Figure 2 and 3 shows the accuracy rates of Korean group and Japanese group. Though from the figures, the rates for minimal pair contrast (LH'-LH) were lower in each group, no significant difference between subject groups was observed (p=0.885).

We found significant main effects for Sequence Length (F (1.073, 21.451) =9.010, p=0.006) and Contrast (F (1.875, 37.499) =11.777, p <0.001). No statistically significant effects were found in interaction between Length and Group (F (1.073, 21.451) = 0.154, p=0.716), Contrast and Group (F (1.875, 37.499) =1.824, p=0.177), Length and Contrast (F (4, 80) = 0.978, p=0.424), and the interaction of all the three factors (F (4, 80) = 1.608, p=0.180).

A Levene Test of Homogeneity of Variances was also conducted in Group as independent factor in each contrast. The significance levels were set to 0.05. No significant difference of variance was shown in each contrast (H'-L vs. L-H': p=0.104, H'-L vs. L-H: p=0.276, L-H' vs. L-H: p=0.644)

Figure 2: Accuracy rates of Korean group.

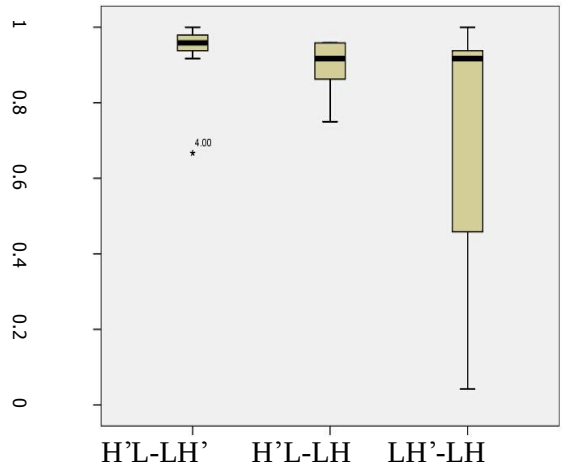
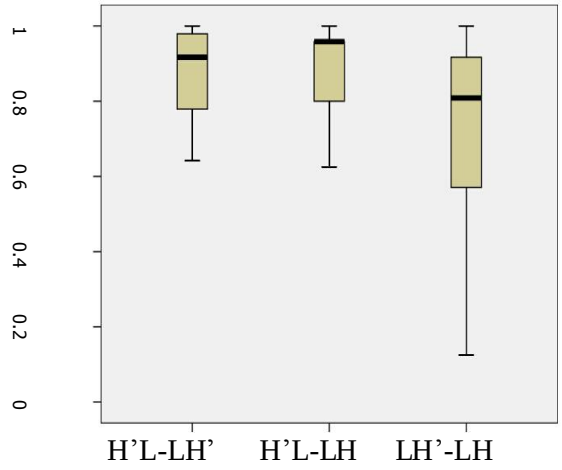


Figure 3: Accuracy rates of Japanese group.



4. Discussion

As shown in the results, the significant main effects in Contrast and Sequence Length showed that there was difference in the average performance between the subjects within the group. However, no statistically significant effect between interactions of two factors or three factors were found, which suggested there was no significant difference between Korean learners of Japanese and native Japanese speakers in perceiving pitch-accent differences. This is different from what we predicted in the beginning.

The model formulating the influence of abstract phonological features in [3-4] failed to account for our observations. However, if we take a look at the language acquisition model proposed by Best [16], the perceptual system assimilates the sound that we hear to the closest category. If the sound we hear is quite close to the already existed sounds in the categories, it will bring difficulties in discriminating them. This indicates the sound will not be assimilated if it is far distinct from any existing categories. If this is the case, it will result in good discrimination [17]. Seoul Korean does not have lexical tone, lexical stress or lexical pitch-accent in word prosody [5]. However, pitch-accent in Japanese is solely pitch based while pitch in Seoul Korean is also crucial in phrasal level. Hence it is hard to tell whether the sound featured pitch-accent is far distinct from any existing categories in the Korean listeners.

Another theoretical model can provide us some insight for our result is the effect of 'acoustic warping' [18-19]. It predicts the perception of certain cues can be enhanced or reduced, depending on the functional values in the native language [18-20]. The main cue for pitch-accent is pitch-based. Though pitch does not function in word-level in Seoul Korean, it plays a crucial role in word segmentation [21]. IP (intonational phrases) and AP (accentual phrases) depend to a large extent on the usage of pitch to distinguish phrasal-level meanings [22-24]. This means, the usage of pitch that functions at phrasal-level in Korean, but functions both at word-level and phrasal-level in Japanese, is enhanced in Korean learners of Japanese. This account still needs to be supported by testing less advanced Korean learners of Japanese.

5. Conclusion

The phonological experiment we applied showed there was no difference between advanced Korean learners of Japanese and native Japanese speakers in perceiving lexical pitch-accent in three different contrasts, including minimal pairs of final accented vs. unaccented. Different from the 'persistent deafness' observed in advanced French learners of Spanish in Dupoux et al. [11], Korean learners of Japanese, a language on a par with French without lexical tonal phenomenon, had no difficulties in perceiving non-native lexical pitch-accent. The present study also served as a tool to examine theoretical acquisition models as discussed above.

Further studies will be carried out in less advanced Korean learners of Japanese. And online parsing of lexical pitch-accent in mental lexicon in Korean learners of Japanese is of interest for future study.

6. Acknowledgement

I would like to thank dr. Manami Hirayama, Marina Arashino, and Hosuk Yun for helping the experiments, and dr. Liquan Liu, Prof. Kubozono and Prof. Kager for their invaluable comments.

7. References

[1] Guion, S. G., Flege, J. E., Akahane-Yamada, R., Pruitt, J. C. 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, 2711–2724

[2] Dupoux, Sebstian-Galles, E., N., Peperkamp, S. 2008. Persistent stress "deafness": the case of French learners of Spanish. *Cognition* 106: 682-706.

[3] Lado, R. 1957. *Linguistics across cultures: Applied linguistics for language teachers*. Ann Arbor, MI: University of Michigan Press.

[4] Brown, C. 1998. The role of the L1 grammar in the L2 acquisition of segmental structure. *Second Language Research*, 14, 136–193.

[5] Jun, S. A. 1993. *The Phonetics and Phonology of Korean Prosody*. Ph.D. diss., Ohio State Univerisity. [Published by Garland Press. New York, 1996]

[6] Haraguchi, S. 1977. *The Tone Pattern of Japanese: An autosegmental Theory of Tonology*. Tokyo: Kaitakusha.

[7] Kim, H. 2010. Korean perception of Japanese Pitch-accents: a case study of native speakers of Seoul Korean. *Japanese-Korean Linguistics* 20. Stanford: CSLI.

[8] Takiguchi, I. 2010. Second language speech perception: pitch cues to Japanese vowel length. Unpublished MA thesis, Kobe University.

[9] Ayusawa, T., Nishinuma, Y., Lee, M. H., Arai, M., Odaka, K., Hoki, N. 1995. Kokusai syakai ni okeru nihongo ni tuite no sogoteki kenkyu kenkyu dai 2 kai kenkyu hookoku-kai yokoo shu. *Sin puro 'nihongo' kenkyu seika hookokusho*, 25-32.

[10] Ayusawa, T., Odaka, K. 1998. Tokyo akusento no kikitori tesuto: 21 gengo gurupu no bogo betsu, seiseki gun betsu seitoritu. Kokusai syakai ni okeru nihongo ni tuite no sogoteki kenkyu kenkyu ronbun shu 1. *Sin puro 'nihongo' sookatsuhan*, 57-71.

[11] Dupoux, E., Peperkamp, S., Sebstian-Galles, N. 2001. A robust method to study stress "deafness".

Journal of the Acoustical Society of America 110: 1606-1618.

[12] Hirayama, T. 1957. Nihongo onchoo no kenkyuu. Tokyo: Meiji Shoin, 1957.

[13] Poser, W. 1984. The Phonetics and Phonology of Tone and Intonation in Japanese. Doctoral dissertation, MIT.

[14] Hayes, B. (2008). *Introductory Phonology*. Wiley.

[15] Van Selst, M, Jolicoeur, P. 1994. A Solution to the Effect of Sample Size on Outlier Elimination, *The Quarterly Journal of Experimental Psychology*, 47A, 631-650.

[16] Best, C. T., McRoberts, G. W., Goodell, E. 2001. Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listeners' native phonological system. *Journal of the Acoustical Society of America*, 109, 775–794.

[17] 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, 229–273. Timonium, MD: York Press

[18] Francis, A. L., Nusbaum, H. C. 2002. Selective attention and the acquisition of new phonetic categories. *Journal of Experiment Psychology: Human Perception and Performan*ce, 28, 349–366.

[19] Jusczyk, P. W. 1997. *The discovery of spoken language*. Cambridge, MA: MIT Press.

[20] Iverson, P., Kuhl, P. K., Akahane-Yamada, R., Diesch, E., Tohkura, Y., Kettermann, A., et al. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition*, 87, B47–B57.

[21] Kim, S., Cho, T. 2009. The use of phrase-level prosodic information in lexical segmentation: Evidence from word-spotting experiments in Korean. *Journal of the Acoustical Society of America*, 125 (5), 3373-3386.

[22] Jun, S.A. 1998. The accentual phrase in the Korean prosodic hierarchy. *Phonology 15*, 189-226.

[23] Jun, S.A. 2000. Korean ToBI labeling conventions. *Journal of Speech Science* 7, 143-169.

[24] Jun, S.A. 2005. Prosodic typology. In *Prosodic Typology: The Phonology of Tone and Phrasing*, Sun-Ah Jun (ed.), 430-458. Oxford: Oxford University Press.