말소리와 음성과학 제 1 권 제 4 호 (2009)
pp. 17~27

Lexical Encoding of L2 Suprasegmentals: Evidence from Korean Learners' Acquisition of Japanese Vowel Length Distinctions

Han, Jeong-Im[1]

ABSTRACT

Despite many studies on the production and perception of L2 phonemes, studies on how such phonemes are encoded lexically remain scarce. The aim of this study is to examine whether L2 learners have a perceptual problem with L2 suprasegmentals which are not present in their L1, or if they are able to perceive but not able to encode them in their lexicon. Specifically, Korean learners were tested to see if they could discriminate the vowel length differences in Japanese at the psychoacoustic level through a simple AX discrimination task. Then, a speeded lexical decision task with high phonetic variability was conducted to see whether they could use such contrasts lexically. The results showed that Korean learners of Japanese have no difficulties in discriminating Japanese vowel length contrast, but they are unable to encode such contrast in their phonological representation, even with long L2 exposure.

Keywords: lexical encoding, Korean, Japanese, vowel length, L2

1. Introduction

Adult learners often experience the difficulties in perceiving and producing the phonemes of a second language (L2), especially when the phonemes are not in their native language. Research on the acquisition of L2 phonological contrasts has focused primarily on explaining their relative difficulty based on the relationship between the native language and the target language. Learners have particular difficulty acquiring a novel L2 contrast when they perceive two separate L2 phonemes as mapping to a single native language phoneme. A frequently cited example of this is the acquisition of English /r/ and /l/ contrast by native speakers of Japanese (Aoyama et al., 2004; Bradlow et al., 1997; Bradlow & Pisoni, 1999; Goto, 1971). These English liquid phonemes are known to be matched by Japanese learners to the Japanese phoneme /r/, although they are not equal in acoustic characteristics. Japanese learners neutralize non-native contrasts in perception initially, but even with significant exposure to the target language, their performance remains significantly below that of native speakers (Lively, Logan & Pisoni, 1993; Takagi & Mann, 1995).

In contrast to the studies on the perception and production of L2 phonemes per se, studies focusing on how such phonemes are encoded in the lexical representation remain scarce. It is not clear whether L2 learners ultimately can establish native-like lexical representations with L2 exposure. There have been two opposing experimental results on the lexical representation of L2 phonemes.

First there is evidence in favor of lexical homophony. Pallier et al. (2001) suggested that even highly fluent L2 learners have identical lexical representation for L2 minimal pair words with a phonemic contrast which is not in their L1. They compared auditory repetition priming in two groups of bilingual speakers of

---

1) Konkuk University, jhan@konkuk.ac.kr
(An earlier version of this paper was presented at the 2nd Cross Language Speech Perception Workshop, Acoustical Society of America at Portland. I would like to thank the audience and also three anonymous reviewers for their valuable comments. This work was supported by the Korea Research Foundation Grant funded by the Korea Government (MOEHRD) (KRF-2007-327-A00635)).

Received: October 25, 2009
Revision received: December 8, 2009
Accepted: December 9, 2009

Spanish and Catalan. Pallier et al. (2001) found that for minimal pairs of Catalan words that differed only in one contrast that was not present in Spanish (e.g. /netə/ and /nɛtə/), Spanish native speakers showed repetition facilitation, while Catalan-dominant bilinguals did not show any repetition effect for these minimal pairs. These results showed that such minimal pairs were processed as homophones by Spanish-dominant bilinguals, and that Spanish bilinguals' lexical representations differ from those of native Catalans. Cutler and Otake (2004) observed similar repetition priming results from Japanese speakers' lexical decision of the English words with /r/ and /l/ minimal pairs, and those from native Dutch speakers' lexical decision of the English /ɛ/ and /æ/ minimal pairs.

More recently, Dupoux et al. (2008) showed that homophony of L2 stress patterns is a lasting processing problem instead of a perceptual problem. Their study started from the observation that L2 learners displayed a very strong impairment to perceive L2 suprasegmentals in tasks using high phonetic variability and memory load, while they showed similar performance with native speakers in a less demanding task such as a single AX discrimination. Focusing on "stress", they raised a question whether L2 learners lack proper phonological representation of such suprasegmentals or they simply lack a metalinguistic representation of contrastive stress. As contrastive stress does not exist in French, and their stress is not marked in the orthography, French learners were thought not to have a metalinguistic representation of Spanish stress. Dupoux et al. conducted a sequence of recall task and a speeded lexical decision task with word-nonword minimal pairs that differed only in the position of stress with three groups of French learners of Spanish grouped according to their general proficiency in Spanish. Their results showed that all three groups of late learners of Spanish had impairment in short-term memory encoding of the stress contrasts, and also in the use of stress to access the lexicon. Based on these results, they argued that French learners of Spanish cannot encode contrastive stress in their phonological representation.

On the other hand, Weber and Cutler (2004) and Cutler et al. (2006) proposed that L2 speakers encode the novel L2 contrast, but they simply do not perceive such contrast in the on-line auditory word identification task. They employed an eye-tracking technology to measure how listeners evaluate incoming auditory input over time. Weber and Cutler (2004) found that Dutch learners of English fixated their visual attention longer and more frequently on pictures of words containing confusable sounds such as /æ/ and /ɛ/ which are not in Dutch phonemes. For example, Dutch speakers fixated longer on a picture of *pencil* when the target word was *panda* than on a less confusable word pair such as *beetle* and *bottle*. However and more strikingly, the listeners' pattern of inappropriate L2 lexical activation was asymmetric, with the target *panda* causing activation of *pencil* but not vice versa. Weber and Cutler (2004) argued that this asymmetry was attributed to Dutch speakers' encoding of both /æ/ and /ɛ/ contrast in their lexicon, and Dutch speakers simply have difficulty perceiving this contrast in perceptual representation. Cutler, Weber and Otake (2006) tried to explain a lexical dominance of a certain phoneme (e.g., /ɛ/ in Dutch) with the phonetic closeness, namely, the L2 category that is phonetically closer to their L1 category emerges as the dominant category. Later, Escudero et al. (2008) put forth an alternative to the phonetic closeness explanation: the influence of orthographic information. They argued that knowledge of contrastive spellings for L2 words can lead learners to encode novel phoneme contrasts in the lexicon which they cannot perceive.

More recently, Hayes-Harb and Masuda (2008) investigated how L2 learners establish their lexical representations that encode a novel L2 phonemic contrast. Specifically they examined the abilities of English learners of Japanese at two different proficiency levels (experienced vs. inexperienced learners) to encode consonant length contrast in Japanese. After participants learned a set of Japanese non-words, including single-geminate contrasts (e.g. *meso-messo*), they completed an auditory word-picture matching task (listening test phase) and a naming task (production test phase). The results of the matching task showed that the experienced learners, but not the inexperienced learners, did not differ significantly from the native Japanese speakers. However, for the results of the production task, performance by the experienced learners was between that of the inexperienced learners and the native Japanese speakers. They interpreted these results as suggesting that inexperienced learners are able to detect the difference between singleton-geminate consonants in the auditory input and encode pairs of newly learned words differing only in consonant length contrastively, but their production abilities neutralize such contrasts. Additionally, one year of Japanese experience significantly improved learners' ability to establish a novel L2 contrast lexically, and thus many of experienced learners were able to accurately encode the feature

[+long], and not just the surface acoustic information, in their lexical representations to accurately differentiate singleton and geminate consonants. As such, the Hayes-Harb and Masuda (2008) study was inconclusive as to whether second language learners are able to store the contrast for Japanese singleton and geminate consonants lexically.

The aim of this paper is to further probe the lexical encoding of L2 suprasegmentals with a different type of a suprasegmental contrast and in a different target language-native language setting from the previous studies. Specifically the question of whether Korean learners have a perceptual problem with the vowel length contrast in Japanese which are not in Korean, or whether they are able to perceive such L2 suprasegmentals but not able to encode them in their phonological representation will be addressed. While vowel length signals a phonemic contrast such as /nu:n/ "snow" vs. /nun/ "eye" in some dialects of Korean, its phonemic function has completely disappeared in the modern standard Seoul dialect (Kim & Han, 1998). In Japanese, on the other hand, vowel and consonant lengths are used phonemically to distinguish words (e.g. *kado* "corner" vs. *kaado* "card" and *ita* "exited" vs. *itta* "said") (Fujisaki et al., 1975; Tajima et al., 2008; Uchida, 1998). Phonologically the length contrast involves the addition of an extra mora, which is marked by different orthographic conventions. The present paper investigates whether Korean learners of Japanese could discriminate the vowel length differences at the psychoacoustic level, utilizing a simple AX discrimination task of vowel length contrast. And then following Dupoux et al. (2008), a speeded lexical decision task with high phonetic variability is conducted to investigate whether Korean learners of Japanese are able to use such vowel length contrasts to access the lexicon.

2. Method

2.1 Materials

2.1.1 Vowel length perception

Twenty Japanese non-word pairs ("length-controlled words") that differed only in vowel length were constructed with a CVCV structure. Ten pairs had a long vowel in the initial syllable (e.g. *gaano* vs. *gano*), and the other ten pairs had a long vowel in the final syllable (e.g. *wamuu* vs. *wamu*). An additional 20 non-word pairs ("segment-controlled words") were constructed for controls where only one segment is different between the two members of a pair (e.g. *tinu* vs. *tibu*). The list of test materials was presented in Appendix A. All the stimuli were recorded by a female Japanese native speaker with Nagoya dialect in a sound-proof room using a Tascam HD-P2 solid-state recorder and a Shure KSM44 microphone. Using the PRAAT speech analysis software package (Boersma & Weenink, 2001), the recorded tokens were digitized at 22,050 Hz and 16 bit resolutions.

2.1.2 Vowel length encoding in the lexicon

Fifty word-nonword pairs that differed only in the vowel length ("length words" vs. "length non-words") were selected as test items (See Appendix B). Thirty two items were bisyllabic, eight of the real words had a long vowel in the initial syllable, another eight a long vowel in the final syllable, and the remaining 16 items no long vowels in the real words. Eighteen word pairs were trisyllabic, three items each of the real words had a long vowel in the initial, medial, and final syllables and the remaining nine words had no long vowels in the real words. The list of test materials is presented in Appendix B. Additional 50 word-nonword pairs that differed only in one phoneme were chosen as controls ("segment words" vs. "segment non-words"). The number of syllables and long vowel position in the word is identical to the test items. All the stimuli were recorded by two Japanese native speakers (1 male and 1 female) in a sound-proof room using a Tascam HD-P2 solid-state recorder and a Shure KSM44 microphone. The female speaker was the same speaker who produced the stimuli words in Experiment 1. The male speaker had a Tokyo dialect. Using the PRAAT speech analysis software package, the recorded tokens were digitized at 22,050 Hz and 16 bit resolutions.

2.2 Participants

Thirty nine native speakers of Korean designated as late learners of Japanese participated in the experiment, with 19 males and 20 females between 19-years and 28-years of age. In addition, 15 native Japanese speakers served as controls (3 males, 12 females). The mean age of the Japanese controls was 27.6 years old, ranging between 20-years-old to 36-years-old. The 39 late learners were further grouped into two groups, beginners and advanced learners. The advanced group of learners (N=19) had received test scores over 720 in JPT or the first level in JLPT, and majorities (11 out of 19) had experiences of living and/or studying in Japan for more than one year. The beginner group (N=20) had no such residence experience in Japan and only three out of 20 learners took the Japanese authorized tests, two of whom had received JPT test scores below 600 and the other learner received the third level in JLPT. The length of learning for each group ranged from 1 year to 8 years. However, it was found that they did not learn Japanese intensively during those years, suggesting that the absolute length of learning cannot be closely correlated with their proficiency. All participants filled out a questionnaire concerning their Japanese language background and then took a short Japanese listening test with 20 questions (excerpted from the JLPT) to check their Japanese phonological proficiency. A vocabulary test was carried out to the beginner group subjects where they were asked to translate Japanese words used in the lexical decision task to Korean. The purpose of the latter test was to confirm that the subjects were familiar to the target words. Only the data for the subjects who wrote down the meaning of the test words more than 90% correct were included for analysis, and thus only the data from 16 beginner group participants out of 20 were used for analysis. <Table 1> summarizes the biographic data and Japanese language background of the two groups of late learners. More detailed information is presented in Appendix C.

Table 1. Biographic data and language background of two groups of Korean late learners of Japanese

	Learners	
	Beginner (N= 16)	Advanced (N= 19)
Mean age	22 years	23.8 years
Mean age of first instruction	17.8 years	16.5 years
Mean length of residence in Japan	0 month	25.5 months
Mean length of Japanese instruction in school	33.1 months	43.5 months
Regularly visits Japanese-speaking countries	12.5%	21%
Regularly speaks Japanese in private life	0%	36.8%
Regularly speaks Japanese in Professional/student life	37.5%	52.6%
Test scores of the listening test (excerpted from JLPT)	9.25/20	13.7/20

2.3 Procedure

Participants were tested individually. Participants were first asked to complete a questionnaire concerning biographical and language background information. They then completed an AX discrimination test. In this procedure, listeners were asked to hit the key on the keyboard representing the same or different tokens after judging whether a newly provided item (X) was physically identical to the given item (A). The subjects responded by pressing one of the two keys arranged horizontally on the number pad on the keyboard: they were asked to hit the key for "1" when they judged the items were identical, and "3" when judging that the two items were different. The measurement of the reaction times (RT) began with the presentation of the target. All subjects were provided a 15-item practice session, which was followed by the test trials. The subjects were informed that the words they would hear would consist of Japanese non-words. For each participant, the order of test word sequence was randomized and the experiment lasted around 6 minutes.

After the AX discrimination, the listeners were provided a short break, before the lexical decision task began. During the lexical decision task, listeners heard the stimuli and were asked to indicate as quickly and accurately as possible if the stimuli presented were real words or not by pressing one of two labeled buttons with their dominant hand. They reacted by pressing on one of the two keys arranged horizontally on the number pad on the keyboard: they were asked to hit the key for "1" when they judged the items were real Japanese words, and "3" when judging that the items were not Japanese words. Before the test phase, 15 items were presented to familiarize themselves. The test lasted for approximately 16 minutes. In both AX discrimination and lexical decision tasks, the interstimulus interval (ISI) was 500ms, and the inter-trial interval was 3000ms, but the subjects were instructed to decide as quickly and accurately as possible. All trials were presented to the listeners through headphones (Sennheiser HD 590) and they were presented in a different random order for each participant, using SuperLab Pro 2.04 software (Cedrus Corporation). After the lexical decision task, each participant took the test to write down the meaning of each word to prohibit the low-level learners from judging the real words as non-words simply because they did not know them (the vocabulary test as described in *Participants*). On average, together the lexical decision task and the vocabulary test lasted around 30 minutes (20 + 10).

3. Results

3.1 Vowel length perception

Mean percent correct discrimination scores and RT of the three groups of listeners were analyzed utilizing two separate repeated measures of ANOVAs for correct scores and RT. The correct scores yielded a significant effect of Word Type (length-controlled vs. segment-controlled) [$F(1, 47)=51.29$, $p<.001$], but there were no significant effect of Group (Beginners vs. advanced learners vs. Japanese natives) [$F(2, 47)=1.20$, $p>.05$], and of the interaction of Group and Word Type [$F(2, 47)=1.42$, $p>.05$]. ANOVAs for RT data to the correct responses revealed that there was a main effect of Word Type [$F(1, 47)=39.10$, $p<.001$], but there was no significant effect of Group [$F(2, 47)=.54$, $p>.05$] and the interaction between Group and Word Type [$F(2, 47)=.78$, $p>.05$]. The mean percent correct discrimination scores and RT to the correct responses for each group are presented in <Figure 1> and <Figure 2> respectively.

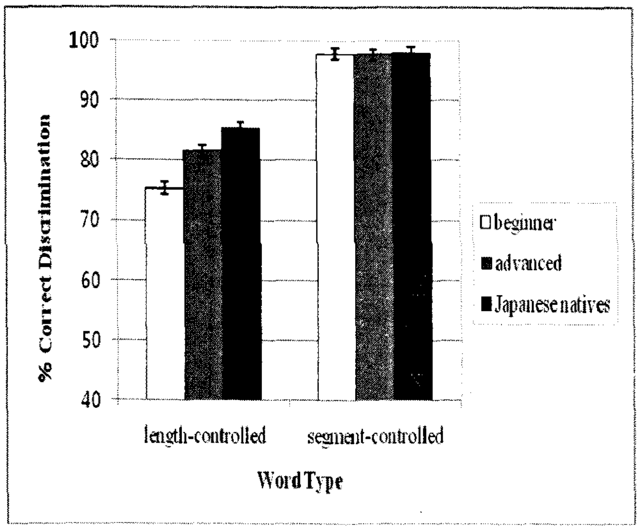

Figure 1. Mean percent correct discrimination scores for three groups

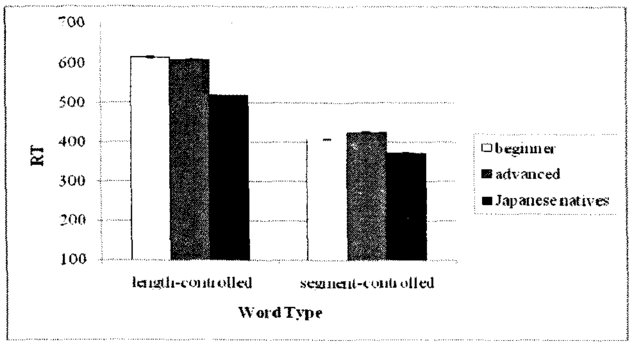

Figure 2. Mean RT for three groups

As in <Figure 1>, all three groups of subjects showed significantly better performance in the perception of segment-controlled word pairs (mean=97.7%) than the length-controlled words (mean=80.7%), suggesting that vowel length contrasts are less discriminated than other segmental contrasts. However and more importantly, there was no significant difference in the discrimination scores among the three groups of subjects. This finding indicated that even the beginner level of Korean learners of Japanese have similar psychoacoustic abilities to more proficient Korean learners and even Japanese natives in perceiving the acoustic differences the long and short vowels show in Japanese.

<Figure 2> indicates that as in the data for correct scores, all three groups of subjects responded to the stimuli significantly faster to the segment-controlled word pairs (mean=400.9 ms) than the length-controlled words (mean=579.7 ms). Similarly to the results of correct scores, Korean learners of Japanese responded to the stimuli at a similar rate as Japanese monolinguals. Again this result revealed that Korean learners of Japanese have no difficulties discriminating the target contrasts.

Putting the results in <Figure 1> and <Figure 2> together, it is suggested that both inexperienced and experienced Korean late learners of Japanese have no impairment in short-term memory encoding of the vowel length contrasts, even though Korean has no such contrasts. However, good performance on the AX discrimination may not be directly associated with evidence that they correctly specified the vowel length contrasts in their lexical representation. Korean learners of Japanese could discriminate the minimal pairs with a vowel length contrast using an ad hoc acoustic strategy. Namely, they were able to detect the difference between short and long vowels in the auditory input, but they simply might differentiate such vowel length contrast based on some phonologically relevant information, which is not necessarily specified as the feature of length (Cutler et al., 2006; Hayes-Harb & Masuda, 2008). Namely, the late learners are able to attend to the acoustic differences between the long and short vowels, but those differences are not necessarily based on the distinctive feature of length. In order to test whether L2 learners are able to use the vowel length contrast in on-line lexical access, Japanese word-nonword minimal pairs differing by a single vowel length or a single phoneme were given to the late learners and they were instructed to make a speeded lexical decision.

3.2 Vowel length encoding in the lexicon

Mean percent correct discrimination scores of the three groups of listeners under four different types of test words (length word, length nonword, segment word, segment nonword) were analyzed utilizing a repeated measures of ANOVA with between-subject factors Group type (beginner vs. advanced vs. Japanese controls) and within-subject factor Word type (length word vs. length nonword vs. segment word vs. segment nonword). The analyses yielded significant main effects of Group type [$F(2, 47)=18.62$, $p<.001$], Word type [$F(3, 141)=42.87$, $p<.001$], as well as an interaction between these two factors [$F(6, 141)=2.93$, $p<.05$]. The results for each group and word type are shown in <Figure 3>.

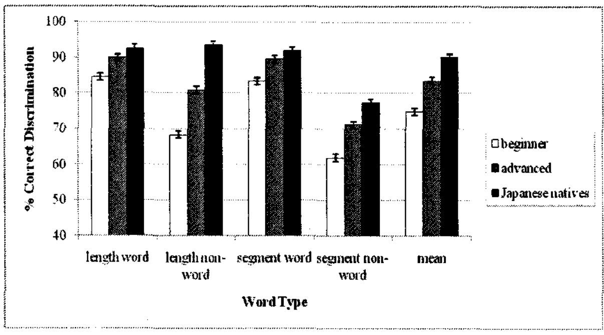

Figure 3. Mean percent correct discrimination scores for three groups

As shown in the rightmost bar graphs in <Figure 3>, there were clear differences in performance among the three groups of subjects: As expected, the discrimination scores of the native listeners were the highest (mean=90.1%) and those of the beginner level of Japanese listeners were the lowest (mean=74.9%). The advanced level of Japanese listeners showed the intermediate scores (mean=83.5%). The post-hoc test (Tukey HSD) showed that there was significant difference among the three levels of learners ($p<.001$ for all pairs). On the other hand, the post-hoc test (Tukey HSD) for Word type factor showed that there was no significant difference between length word and segment word ($p>.05$), while all other pairs yielded significant differences ($p<.001$).

The significant Word type x Group type interaction was explored through tests of simple main effects, which revealed main effects of Word types for beginner level of learners [$F(1, 47)=19.35$, $p<.001$], and less clearly for advanced level of learners [$F(1, 47)=6.76$, $p=.012$], but there was no such significant effect for Japanese controls [$F(1, 47)=.07$, $p>.05$]. These results indicate that the Korean learners of Japanese showed more errors in the nonwords than real words, while the Japanese controls made very few errors overall without such difference between them. Korean learners of Japanese tended to misidentify the non-words as words much more often than they rejected real words, while the Japanese controls showed similar results between real words and non-words. The late learners' pattern of results obtained in the segment words was somewhat different from those in the length words. The difference in the correct scores between the segment words and nonwords was shown to be significant for all subject groups [$F(1, 47)=33.87$, $p<.001$ for beginner level of learners; $F(1, 47)=29.52$, $p<.001$ for advanced level of learners; $F(1, 47)=14.47$, $p<.001$ for Japanese controls]. Given the relatively low scores of segment nonwords for Japanese controls, no meaningful pattern can be derived from the results of segment words and nonwords.

<Figure 4> indicates the mean reaction times (RT) of the three groups of listeners under two different types of test words. A repeated measures of ANOVAs on RT revealed only marginally significant effect of Group type [$F(2, 47)=3.06$, $p=.056$], but there was significant effect of Word type [$F(3, 141)=52.48$, $p<.001$], and the interaction between Group type and Word type [$F(6, 141)=3.32$, $p<.05$].

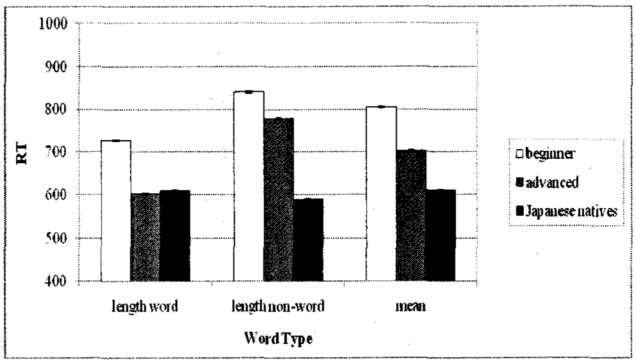

Figure 4. Mean RT for three groups

Though it was shown to be only marginal difference, the rightmost graph in <Figure 4> clearly showed that overall the beginner levels of Korean learners (mean=805.2 ms) responded to the stimuli slower than the advanced levels of Korean learners (mean=703.1 ms). Neither group of Korean learners was faster than the Japanese controls (mean=609.2 ms).

The results of word type factor were analyzed in more detail in terms of running ANOVAs on the RT difference only for length words, because no meaningful pattern was derived from the results of segment words. First in the length words, the difference in RT between the words and nonwords was shown to be significant for Korean learners of Japanese [$F(1, 47)=17.87$, $p<.001$ for beginner level of learners; $F(1, 47)=19.17$, $p<.001$ for advanced level of learners], but not for the Japanese native speakers [$F(1, 47)=.07$, $p>.05$]. These results indicate that both beginner and advanced level of late learners responded to the length words much faster than to the length nonwords, suggesting that they tended to identify length nonwords as real words. However, Japanese natives did not show such difference between real words and nonwords.

These results seem to support the findings of Dupoux et al. (2008) where French late learners of Spanish made more errors for real words than for non-words, 24 % vs. 58% in the test condition. The present results as well as the previous results suggest that in a lexical decision task, late learners had little difficulty in the perception of vowel length of real words, because those words were familiar to the late learners. However, they seemed to identify the nonwords, which were different from real words only in terms of vowel length, as real words. Even the advanced level of learners showed significant difference in the perception scores and RT from the native speakers.

4. Discussion

The present study investigated whether L2 learners are able to encode a novel phonemic contrast in the lexicon. Korean has no vowel length contrast in its phonological representation, which led us to expect that Korean learners of Japanese not only have difficulties in discriminating Japanese vowel length contrast, but also be unable to encode such contrast in their lexical representation.

In Experiment 1, the beginner and advanced levels of Korean learners were asked to discriminate the Japanese word-nonword pairs that differed only in vowel length and showed that both groups of Korean learners as well as native Japanese speakers discriminated the minimal pairs quite successfully in a simple AX discrimination task. The results of Experiment 1 indicate that Korean learners of Japanese do not have any difficulties in perceiving the vowel length contrast. The important point to be made is that L2 learners are not totally deaf to the suprasegmentals in L2 which are not in their L1, and they are sensitive to those distinctions. If they are given sufficient time and resources, they are capable of attending to the acoustic differences between the long and short vowels similar to native speakers. However, it is not clear whether L2 learners in the present study distinguish the vowel length contrasts in L2, based on the feature of length or relevant acoustic information.

In Experiment 2, we examined the use of vowel length in lexical representation by means of a lexical decision task, which was more demanding in terms of memory load and talker variability. Korean learners showed poorer performance than the native Japanese speakers in terms of both the correct scores and RT; there was also significant difference between the two groups of Korean learners. Korean learners of Japanese perceived the vowel length contrast in a lexical decision task for real words with relatively high frequency; however, they frequently misidentified the non-words as real words. Thus the length distinction that is pertinent to the encoding of words in Japanese barely seems to be available to Korean learners of Japanese.

The results of Experiment 1 are partly consistent with those of Hayes-Harb and Masuda (2008) where English learners of Japanese were able to discriminate the minimal pairs with a singleton-geminate consonant contrast difference (*meso* vs. *messo*) without any significant difference from Japanese natives. At the level of perception, both inexperienced and experienced subjects could detect the difference between short and long segments in the auditory input. However, the present results are at odds with that of Dupoux et al. (2008) where French learners of Spanish, whether they are inexperienced or experienced, showed the same impairment in short-term memory encoding of stress contrasts. Specifically in a sequence recall test, French learners made around three times more errors than Spanish natives and more importantly, there was no improvement even after a significant exposure in the perception of contrastive stress. On the other hand, the results of Experiment 2 are in good agreement with the results from Dupoux et al. (2008) study where L2 learners made more errors in the lexical encoding of the L2 stress contrasts than the native speakers. And they showed more errors for nonwords than for words, irrespective of their level of proficiency, suggesting that they predominantly classified the nonwords as words. The present results are again partly consistent with those of Hayes-Harb and Masuda (2008) in that in both studies the L2 learners showed statistically significant improvement in the lexical encoding of the suprasegmentals in L2. However, only in the present study was there a large gap in the performance between the advanced L2 learners and the natives.

The discrepancies among the results of the studies including the present study can be explained based on a number of factors such as interference from existing suprasegmental contrasts in the native language of subjects, lexicality of test words, frequency, task differences, etc., but crucially seem to be related to the target contrast and/or the processing load of the task. First, the present study as well as Hayes-Harb and Masuda (2008) employed a recording of a single speaker, while six speakers (3 males and 3 females) were used in Dupoux et al. (2008). Indeed, the introduction of memory load and talker variability could lead the subjects to undergo a very strong impairment to perceive stress contrasts. Allen and Miller (2004) observed that even L2 learners encounter a number of talkers whose speech exhibits idiosyncratic phonetic characteristics, and thus listeners pay attention to the acoustic cues identifing a particular talker's voice. Crucially these idiosyncratic aspects of pronunciation are utilized in lexical processing as well. Bradlow and Pisoni (1999), Nygaard, Sommers and Pisoni (1994) demonstrated that listeners performed better in word recognition with familiar talkers than unfamiliar talkers. Similarly, it was shown that listeners showed improved performance in the word recognition of novel words after a perceptual training with low-intelligibility speech (Davis et al., 2005; Greenspan, Nusbaum & Pisoni, 1998; Schwab, Nusbaum & Pisoni, 1985). All these results imply that listeners could employ the minute acoustic details of talkers in the lexical encoding of words. As briefly mentioned earlier, L2 learners could employ relevant acoustic information to distinguish the L2 suprasegments, but in more demanding tasks, they might have difficulties in using such information to distinguish the L2 contrast lexically.

There also might be degree of difficulty in the perception of suprasegmentals. Not all suprasegmentals are equally difficult to process. The present study and Hayes-Harb and Masuda (2008) study targeted the length of contrasts, namely vowel length contrast in the present study, and consonant length contrast in Hayes-Harb and Masuda's study. On the other hand, Dupoux et al. (2008) examined the L2 learners' perception of lexical stress. Each type of suprasegmentals involves different types of acoustic cues. For example, stress is associated with duration, intensity, and pitch change, while length contrast seems to be associated with rather smaller numbers of acoustic cues. Stress also involves more global acoustic cues than the length contrast. These cue differences might cause discrepancies in the perception of L2 sound pairs with those contrasts.

Despite all those methodological differences, the finding that L2 learners had difficulty in encoding the suprasegmentals lexically in more demanding tasks suggests that such suprasegmentals are not robust for non-natives.

Taken together the results of Experiment 1 and Experiment 2 clearly indicate that Korean late learners of Japanese could use acoustic cues for vowel length in an acoustic representation, but not in an abstract phonological representation. The findings of the present study run counter to the argument made by Weber and Cutler (2004) and Cutler et al. (2006) that L2 speakers encode the novel L2 contrast, but they simply do not perceive such contrast in the on-line auditory word identification task. As to the question of how novel L2 contrasts are lexically encoded by learners, which they cannot perceive, Cutler and Otake (2006) proposed that the L2 category that is phonetically closer to the L1 category emerges as the dominant category, and Escudero et al. (2008) argued that knowledge of contrastive spellings for L2 words can lead learners to encode novel phoneme contrasts in the lexicon, which they cannot perceive. Interestingly, though stress is not marked in orthography of the target language and thus it is not taught explicitly in school, both consonant and vowel length in Japanese are marked in a written form, with the use of an additional letter for Japanese, which could help learners to establish the distinct representations of long and short segments. However, the present results showed L2 learners' impairment in encoding lexical length difference in their L2 lexicon. Again other methodological effects such as the processing load might overrule the effects of orthography, and further study is necessary to clarify this issue.

L2 learners' use of acoustic cues for vowel length in an acoustic representation, but not in a lexical representation could be evidence for an intermediate stage between having a discrimination profile indicative of sensitivity to a phonemic contrast, and being able to exploit the contrast lexically. Namely, L2 learners' ability to discriminate the novel contrast precedes their ability to represent the contrast lexically. The intermediate stage for the phonemic vs. lexical perception was already attested in studies of first language acquisition (Pater et al., 2004; Stager & Werker, 1997; Werker & Stager, 2000). It is widely known that across the first year of life, infants progress from an initial stage of sensitivity to considerable phonetic detail to a selective focus on the phonetic variability that occurs in the native language (for a review see Jusczyk, 1997). However, another important development that occurs during the first year of life is infants' ability to recognize words (Halle & Boysson-Bardies, 1994; Jusczyk & Aslin, 1995; Myers, Jusczyk, Kemler-Nelson, Charles-Luce, Woodward & Hirsh-Pasek, 1996). Werker and Stager

(2000) suggested that infants gradually develop the ability to pull out actual words from the speech stream, distinguishing one word from another, after they attain the phonemes in their native language. When the infants are first learning the association between a novel word and a novel object, they are not able to use minimally contrastive phonetic detail, even though they are quite capable of discriminating the detail. However, infants gradually develop the skills to pick up the phonetic details necessary for recognizing the words. Once children are listening to distinguish meaningful lexical items, the representation used on comprehension may contain less phonetic detail (e.g. Kay-Raining Bird & Chapman, 1996)

The results of the present study as well as previous studies suggest that even L2 learners show such developmental changes from the speech perception to word processing. Most L2 learners in the present study and previous studies seem to be at the point in development where they are able to detect and discriminate the L2 suprasegmentals, but still have some difficulties to use those phonetic properties to pull out actual words. They seem to increasingly construct the phonological representation of L2, but without attaining the level of native speakers. How much of L2 experience is enough to establish native-like lexical representation is not clear at this moment, which needs a further study.[2]

References

Allen, J. S. & Miller, J. L. (2004). "Listener sensitivity to individual talker differences in voice-onset-time", *Journal of the Acoustical Society of America,* Vol. 115, pp. 3171-3183.

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 /l/ and /r/", *Journal of Phonetics,* Vol. 32, pp. 233-50.

Boersma, P. & Weenink, D. (2001). "Praat: Doing phonetics by computer", *Glot International,* Vol. 5(9/10), pp. 341-345.

Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R. & Tohkura, Y. (1997). "Training Japanese listeners to identify English /r/ and /l/: volume IV: Some effects of perceptual learning on speech production", *Journal of the Acoustical Society of America,* 101, pp. 2299-2310.

Bradlow, A. R. & Pisoni, D. B. (1999). "Recognition of spoken words by native and nonnative listeners: Talker-, listener-, and item-related factors", *Journal of the Acoustical Society of America,* Vol. 106, pp. 2074-2085.

Cutler, A. & Otake, T. (2004). "Pseudo-homophony in non-native listening", Unpublished paper presented at the Acoustical Society of America, New York.

Cutler, A., Weber, A. & Otake, T. (2006). "Asymmetric mapping from phonetic to lexical representations in second-language listening", *Journal of Phonetics,* Vol. 34, pp. 269-284.

Davis, M. H., Johnsrude, I. S., Hervais-Adelman, A., Taylor, K. & McGettigan, C. (2005). "Lexical information drives perceptual learning of distorted speech: Evidence from the comprehension of noise-vocoded sentences", *Journal of Experimental Psychology: General,* Vol. 134, pp. 222-241.

Doupoux, E., Sebastian-Galles, N., Navarrete, E., & Peperkamp, S. (2008). "Persistent stress 'deafness': The case of French learners of Spanish", *Cognition,* Vol. 106, No.2, pp. 682-706.

Escudero, P., Hayes-Harb, R. & Mitterer, H. (2008). "Novel second-language words and asymmetric lexical access", *Journal of Phonetics,* Vol. 36, pp. 345-360.

Fujisaki, H., Nakamura, K. & Imoto, T. (1975). "Auditory perception of duration of speech and non-speech stimuli", In G. Fant & M. A. A. Tatham (Eds.), Auditory analysis and perception of speech (pp. 197-219). Academic: London.

Goto, H. (1971). "Auditory perception by normal Japanese adults of the sounds 'l' and 'r' ", *Neuropsychologia,* Vol. 9, 317-23.

Greenspan, S. L., Nusbaum, H. & Pisoni, D. B. (1998). "Perceptual learning of synthetic speech produced by rule", *Journal of Experimental Psychology: Learning, Memory, and Cognition,* Vol. 14, pp. 421-433.

Halles, P. A. & Boysson-Bardies, B. de (1996). "The format of representation of recognized words in infants' early receptive lexicon", *Infant Behavior and Develop*ment, Vol. 19, pp. 463-481.

Hayes-Harb, R. & Masuda, K. (2008). "Development of the ability to lexically encode novel second language phonemic contrasts", *Second Language Research,* Vol. 24, No. 1, pp. 5-33.

Jusczyk, P. (1997). The discovery of spoken language. Cambridge, MA: MIT Press.

Jusczyk, P. & Aslin, R. N. (1995). "Infants' detection of the sound patterns of words in fluent speech", *Cognitive Psychology,* Vol. 29, pp. 1-23.

Kay-Raining Bird, E. & Chapman, R. S. (1996). "Partial representations and phonological selectivity", Unpublished manuscript.

Kim, H.-S. & Han, J.-I. (1998). "Vowel length in modern Korean: An acoustic analysis", In B.-S. Park & J. H. Yoon (Eds.), Proceedings of the 11th International Conference on Korean (pp. 412-418), Seoul, Korea: Hanshin.

Lively, S. E., Logan, J. S. & Pisoni, D. B. (1993). "Training Japanese listeners to identify English /r/ and /l/ II: The role of phonetic environment and talker variability in learning new perceptual categories", *Journal of the Acoustical Society of America,* Vol. 94, pp. 1242-1255.

Myers, J., Jusczyk, P., Kemler-Nelson, D. G., Charles-Luce, J., Woodward, A. & Hirsh-Pasek, K. (1996). "Infants' sensitivity to word boundaries in fluent speech", *Journal of Child Language,* Vol. 23, pp. 1-30.

Nygaard, L. C., Sommers, M. S. & Pisoni, D. B. (1994). "Speech perception as a talker-contingent process", *Psychological Science,* Vol. 5, pp. 42-46.

[2] A reviewer pointed out that in Japanese pitch variation as well as vowel and consonant length distinctions are phonemic and even more important suprasegmental properties, and thus should be tested for a clearer picture of the exical encoding of L2 suprasegmentals.

Pallier, C., Colome, A. & Sebastian-Galles, N. (2001). "The influence of native-language phonology on lexical access: Exemplar-based vs. abstract lexical entries", *Psychological Science,* Vol. 12, pp. 445-449.
Pater, J., Stager, C. & Werker, J. (2004). "The perceptual acquisition of phonological contrasts", *Language, 80*, 384-402.
Schwab, E. C., Nusbaum, H. C. & Pisoni, D. B. (1985). "Some effects of training on the perception of synthetic speech", *Human Factors,* Vol. 27, pp. 395-408.
Stager, C. & Werker J. (1997). "Infants listen for more phonetic detail in speech perception than in word-learning tasks", *Nature,* 388, 6640, pp. 381-382.
Takagi, N. & Mann, V. (1995). "The limits of extended naturalistic exposure on the perceptual mastery of English /r/ and /l/ by adult Japanese learners of English", *Applied Psycholinguistics,* Vol. 16, pp. 379-405.
Tajima, K., Kato, H., Rothwell, A., Akahane-Yamada, R. & Munhall, K. G. (2008). "Training English listeners to perceive phonemic length contrasts in Japanese", *Journal of the Acoustical Society of America,* Vol. 123, No.1, pp. 397-413.
Toda, T. (2003). "Second language speech perception and production: Acquisition of phonological contrasts in Japanese", Lanham, MD: University Press of America.
Uchida, T. (1998). "Categorical perception of relatively steady-state speech sound duration in Japanese moraic phoneme", *Journal of Phonetic Society of Japan,* Vol. 2, pp. 71-86 [in Japanese]
Weber, A. & Cutler, A. (2004). "Lexical competition in non-native spoken-word recognition", *Journal of Memory and Language,* Vol. 50, pp. 1-25.
Werker, J. & Stager, C. (2000). "Developmental changes in infant speech perception and early word learning: Is there a link? In M. Broe, & J. Pierrehumbert (Eds.)", *Papers in Laboratory Phonology V: Acquisition and the Lexicon* (pp. 181-193), Cambridge University Press.

Appendix A

Items for auditory listening (Experiment 1)

Length-controlled items where the first syllable has a long vowel

gaano-gano, saako-sako, deeke-deke, hootu-hotu, tiinu-tinu, haahi-hahi, haaro-haro, waami-wami, reeha-reha, reema-rema

Length-controlled items where the second syllable has a long vowel

wamuu-wamu, gahee-gahe, desaa-desa, wanuu-wanu, deroo-dero, hanuu-hanu, makoo-mako, rumee-rume, tusoo-tuso, remii-remi

Segment-controlled items where a single consonant is different

tinu-tibu, ruya-muya, gaahi-saahi, hotu-dotu, dinu-dibu, hahi-habi, haaro-saaro, remii-rezii, reeha-beeha, wanuu-yanuu

Segment-controlled items where a single vowel is different

gano-ganu, reeha-riiha, sako-suko, dene-done, dinu-donu, hahi-hehi, haaro-heero, remii-romii, gahe-gohe, ruuya-riiya

Appendix B

Items for lexical decision (Experiment 2)

Length items (bisyllabic)

doogu, tuuka, doozo, byooki, ryoori, puuru, geemu, keeki

doyoo, tugoo, tokee, kinoo, sumoo, guree, menyuu, gitaa

heya, hana, tuma, kami, sohu, gogo, mata, hati

kasa, doko, neko, dame, yoru, yuki, hito, kuzi

Length items (trisyllabic)

teeburu, Beekoku, syuumatu, sukosi, tegami, donata

hikooki, depaato, sapooto, hazime, deguti, hitori

gakusee, kawaii, yasasii, musuko, sigoto, sasimi

Segment items (real words)

ame, anata, bangoo, benri, benkyoo, dame, eegakan, soko, dore, ikura, usiro, kaisya, kisya, kaban, yahari, mae, zutoo, koosaten, kuruma, midori, ototoi, nooto, obaasan, okosan, yukkuri

Segment items (non-words)

ade, amata, bangee, penri, bekyoo, dane, egi, sogo, tore, igaka, tyoto, sitori, muuzi, yukki, yaharu, gazi, sutoo, koosatin, buruma, omoti, otyi, tooto, ozaasan, ekosan, yukuri

Appendix C

Biographical information of participants

1) Advanced group of listeners

listener	sex	AFI	LOR	LOL	SOLT (/ 20)	SOAT
PSL	M	13	96	96	12	JPT 735
KHM	F	18	12	36	13	JLPT 1st
APK	M	17	None	36	10	JLPT 1st
CSH	F	4	60	24	16	JLPT 1st
KJE	F	18	12	36	14	JLPT 1st
LHP	F	17	12	84	16	JLPT 1st
CSI	M	15	None	12	15	JLPT 1st
JPK	M	23	12	12	12	JLPT 1st
LDH	F	13	None	48	11	JLPT 1st
KRY	F	17	12	36	14	JPT 750

LJY	M	16	24	60	17	JPT 965
LDW	M	25	None	48	13	JLPT 1st
LMY	F	18	None	48	17	JLPT 1st
PJE	F	18	24	24	13	JLPT 1st
KSP	M	20	12	48	13	JLPT 1st
SSY	M	15	None	30	12	JPT 775
KJS	M	18	12	53	14	JLPT 1st
NKY	M	17	None	48	16	JLPT 1st
JEJ	F	12	None	48	13	JLPT 1st

(AFI=age of first instruction, LOR=length of residence in Japan (months), LOL=length of learning in schools (months), SOLT=scores of the listening test (out of 20), SOAT=scores of the authorized tests)

2) Beginner group of listeners

listener	sex	AFI	LOR (months)	LOL (months)	SOLT (out of 20)	SOAT
NPS	M	19	None	12	4	
KMP	M	19	None	20	8	
MJY	M	25	None	12	3	
LNH	F	15	None	60	10	
LAR	F	17	None	36	9	
KHR	F	16	None	36	14	JPT 595
SSY	F	17	None	24	12	
STK	M	17	None	60	10	
KPY	F	18	None	48	11	
YSW	M	20	None	12	10	
KDN	F	16	None	24	10	
KYK	F	15	None	48	12	JLPT 4th
JYH	F	18	None	36	6	
KSJ	F	19	None	36	11	
YSH	M	18	None	24	13	JLPT 3rd
JKS	M	16	None	42	5	

(AFI=age of first instruction, LOR=length of residence in Japan (months), LOL=length of learning in schools (months), SOLT=scores of the listening test (out of 20), SOAT=scores of the authorized tests)