PERCEPTION OF KOREAN BACK VOWELS BY AUSTRALIAN ENGLISH AND JAPANESE SPEAKING ADULT LEARNERS

Chong-Woon Kim and John C. L. Ingram
Griffith University & The University of Queensland

ABSTRACT - The results of this study indicate that the nature of perceptual mapping between target L2 vowels and their nearest matching L1 counterparts is crucially related to the ways in which L2 sounds are accommodated by the learners with different L1 backgrounds at different stages of L2 learning.

INTRODUCTION

This paper examines the perception of five naturally produced Korean vowels (i.e. /a/, /ʌ/, /o/, /u/ and /ɯ/) by Australian English (AE) and Japanese (J) speaking adult learners of Korean, each at two levels of proficiency in Korean (beginners, advanced). Subjects were tested on multiple exemplars of the contrasting vowels produced in a [hVda] carrier frame by three native Korean speakers.

The aims of the study are: a) to investigate effects of the language background (learner's L1 vowels) on the acquisition of vowel contrasts in second language learning, b) to investigate the relationship between identification and discrimination in L2 vowel perception, and c) to determine the effects of perceived similarities between the to-be-acquired L2 vowels and their nearest native language (L1) targets or prototypes.

Casual classroom observation indicates that students with various L1 backgrounds have more difficulty with four of these Korean vowels (/ʌ/, /o/, /u/ and /ɯ/) than others in terms of production and perception. One of the practical objectives of the present study with direct pedagogic implications is to describe more clearly and systematically the nature of the difficulty involved in the learning of these vowels by English and Japanese learners of Korean.

In order to address the above objectives, we require data from three different types of experiment. We require information on how well learners categorize or identify the target vowels in L2 (experiment 1: Identification task). Identification and patterns of misidentification reflect the primary performance of learners accommodating to a new sound system. However, most theories of second language sound accommodation require discrimination data to test their predictions. The relationship between identification and discrimination as perceptual tasks has long been an issue of interest (Stevens et al. 1969). Hence, an AXB discrimination task (experiment 2) on the same set of L2 tokens was undertaken. But, as Best's (1995) and other's models of L2 vowel acquisition make clear, the nature of the perceptual mapping between the target L2 vowels and their matching L1 counterparts is crucial, at least in initial stages of L2 accommodation. Consequently, a third experiment was undertaken which involved obtaining perceived similarity ratings of the L2 vowels in relation to native or L1 targets, and prototypicality ratings for the L2 vowels as L1 targets.

METHOD

The subjects were Australian and Japanese learners of Korean enrolled at Griffith University. Each L1 group was divided into two sub-groups according to their proficiency level in Korean, forming four subject groups (i.e. AE1, AE2, J1 and J2). Both the AE1 and J1 group consist of undergraduate students with no in-country experience while AE2 and J2 subjects have minimum of one and half year in-country training. None of the subjects were reported to have any hearing problems. The number of subjects participating in each of the three experiments is given in the table below (table 1). Experiment 3 has the largest number of subjects. Experiment 1 and 2 contains a subset of the subjects who participated in experiment

3. The average age of the subjects was 22.5.

Group	AE 1			AE2			J1			J2		
Experiment	1	2	3	1	2	3	1	2	3	1	2	3
Male	6	6	14	3	3	3	5	5	7	0	0	0
Female	16	20	21	3	3	3	5	5	6	5	5	5

Table 1. Number of AE and J subjects for each experiment

A computer-based delivery system was used for all three experiments. Each experiment contained a series of familiarization trials. All stimuli were delivered binaurally through a headphones connected to each computer. The main purpose of the experiment 1 (identification) was to elicit listener categorization of stimuli with respect to L2 phonological categories. The experiment contained 60 trials (5 vowels x 3 speakers x 2 tokens x 2 times repeated) which were arranged in random order. Identical stimulus was repeated three times in each trial and the inter-stimulus interval (ISI) was 0.5 second. Subjects made their responses by moving the cursor to a chosen L2 keyword and clicking the mouse. The subjects controlled the interval between trials.

Experiment 2 (discrimination) was designed to map systematically how the L2 sounds are discriminated from one another by the two L1 groups. Because of less demand on auditory memory (Polka, 1995), an AXB procedure was utilized in which A and B are always different and the middle token is identical (phonemically not physically) to either A or B. A total of 60 trials were generated (i.e. 10 possible parings from 5 vowels x 3 speakers x repeated 2 times). In order to invoke phonemic not acoustic coding, a relatively longer ISI of 2 seconds was used. Subjects were instructed to mouse click either A or B on the screen after hearing a sequence of three tokens.

The objective of experiment 3 (categorization by L1 keyword and rating) was to ascertain how the L2 sounds are related perceptually to the listeners' L1 categories. The experiment involved two tasks: 1) choosing a native language key word that contains the vowel most similar to the stimulus L2 vowel, 2) giving a prototypicality rating for the chosen L1 vowel using a 7 point rating scale where 7 represents 'identical' and 1 represents 'quite different'. The choice in each case was made by a mouse click. Two different versions of the experiment were developed for each language group using respective native orthography. The main experiment contains 60 trials (5 vowels x 3 speakers x 2 tokens x 2 times repeated) which were arranged in random order. Identical stimulus was repeated three times in each trial and the inter-stimulus interval (ISI) was 0.5 second. The inter-trial interval (ITI) was 4 seconds and the beginning of a new trial was signaled by a short chime-bell sound.

RESULTS

Table 2 summarizes the results of the experiment 1. The table clearly shows that K /ʌ/, /o/, /u/ and /ɯ/ vowels are difficult ones to identify for all subjects, especially when one considers the identification scores for K /a/. Overall, AE groups performed better than J groups and the cross-sectional difference is more evident than the cross language difference.

K. vowel	AE1	AE2	J1	J2
/a/	11.8 (0.4) 98.5%	12.0 (0.0) 100.0%	11.5 (1.3) 95.8%	12.0 (0.0) 100.0%
/ʌ/	5.8 (1.8) 48.0%	8.0 (2.5) 66.7%	6.2 (1.8) 51.7%	7.0 (1.4) 58.3%
/o/	6.6 (4.1) 55.3%	9.5 (1.8) 79.2%	4.3 (3.1) 35.8%	6.2 (3.0) 51.7%
/u/	7.9 (3.3) 65.5%	10.8 (1.2) 90.3%	5.8 (3.8) 48.3%	9.4 (2.6) 78.3%
/ɯ/	6.5 (2.9) 53.8%	8.8 (2.5) 73.6%	8.4 (2.4) 70.0%	8.4 (2.3) 70.0%
mean	7.7 (3.5) 64.2%	9.8 (2.2) 81.9%	7.2 (3.6) 60.3%	8.6 (2.8) 71.7%

Table 2. Identification results (mean, standard deviation, mean average percent correct)

To examine the effect of L1, proficiency level and vowels on the identification performance, ANOVA was applied to the data. The ANOVA results revealed that that the effect of L1 is not significant for the overall pattern of identification performance, confirming the initial observation that all subject groups have difficulty in the identification of these Korean vowels except for /a/. The effect of proficiency level, on the other hand, was significant indicating that there is a substantial difference in the identification performance between the more proficient and the less proficient subject groups regardless of L1. The type of vowel being tested was also significantly related to the variation in the identification performance. In order to rule out the possible statistical bias caused by the exceptionally good performance obtained for the Korean /a/, ANOVA was reapplied without the /a/ vowel and the type of vowel remained a significant factor. Both ANOVA results showed that the interaction between L1 and Vowel factor is significant. This indicates that depending on the vowel being tested there is a language specific influence in the identification performance by different L1 groups.

The results of the experiment 2 (discrimination) are summarized in table 3. As one would expect, the score for correct discrimination is much higher than the score from the identification experiment. Subject groups as a whole, the percentage correct for identification is 69.5% and discrimination is 94.6%. AE group overall performed better than the J group. Those cases where each vowel is separated clearly from the other in terms of height and backness, in general, got very high correct discrimination score regardless of L1 and proficiency level. The discrimination performance for the pairs containing the four difficult vowels was substantially poorer.

K. vowels	AE1	AE2	J1	J2
/a//ʌ/	5.7 (0.5) 97.5%	6.0 (0.0) 100.0%	5.5 (0.7) 91.7%	6.0 (0.0) 100.0%
/a//o/	5.7 (0.5) 97.5%	6.0 (0.0) 100.0%	5.7 (0.7) 95.0%	6.0 (0.0) 100.0%
/a//u/	5.8 (0.5) 96.8%	5.8 (0.4) 97.2%	5.6 (0.7) 93.3%	6.0 (0.0) 100.0%
/a//ɯ/	5.8 (0.4) 96.8%	6.0 (0.0) 100.0%	6.0 (0.0) 100.0%	6.0 (0.0) 100.0%
/ʌ//o/	5.4 (0.9) 89.2%	5.8 (0.4) 97.2%	5.2 (1.3) 86.7%	5.6 (0.6) 93.3%
/ʌ//u/	5.9 (0.3) 98.0%	5.8 (0.4) 97.2%	5.4 (0.8) 90.0%	5.6 (0.6) 93.3%
/ʌ//ɯ/	5.9 (0.4) 98.0%	5.7 (0.5) 94.5%	5.5 (0.7) 91.7%	6.0 (0.0) 100.0%
/o//u/	5.7 (0.8) 94.2%	6.0 (0.0) 100.0%	5.4 (0.8) 90.0%	5.4 (0.6) 90.0%
/o//ɯ/	5.8 (0.5) 96.2%	6.0 (0.0) 100.0%	5.6 (0.7) 93.3%	5.6 (0.6) 93.3%
/u//ɯ/	5.4 (0.6) 89.7%	5.3 (0.5) 88.8%	4.3 (1.1) 71.7%	4.8 (1.3) 80.0%
mean	5.7 (0.6) 95.3%	5.9 (0.4) 97.5%	5.4 (0.9) 90.3%	5.7 (0.6) 95.0%

Table 3. Discrimination results (mean, standard deviation, mean average percent correct)

ANOVA was again applied to the data to examine the effect of L1, proficiency level and vowel contrasts on the discrimination performance. Unlike the results from the identification experiment, proficiency level did not have a significant effect on the discrimination performance. This may be predictable from the fact that many of the vowel pairs contain the well-identified vowel /a/ as well as the possible use of auditory-acoustic coding that may have been employed by some subjects. Significant main and interaction effects were obtained for the factors of L1 and the Vowel Contrast. The interaction effect revealed differences between the Japanese and English learners in their discrimination of some target vowel contrasts.

DISCUSSION

In order to understand the basis of the Vowel by L1 interaction in experiment 1 (identification task) and the Vowel contrast by L1 interaction in experiment 2 (discrimination), we need to look at the results of experiment 3 (categorization by L1 keyword and rating) given in table 4. In table 4 the vowel choice and ratings score were combined to generate a composite similarity index score. Most noticeable from table 4 is the substantially higher composite similarity index score given to K /a/ vowel by all groups indicating that this vowel is perceived to be very similar to a native vowel category for both AE and J groups. This high level of similarity clearly relates to the very high identification score obtained for K /a/ vowel by all subjects. It also explains why AE and J subjects performed exceptionally well when Korean /a/ was contrasted

against the four other vowels in discrimination experiment. If one can clearly identify one vowel in AXB task, one would be highly accurate in discrimination regardless of how well or poorly the other vowel is categorized.

K vowel	AE1 response	AE2 response	J1 response	J2 response
/a/	/a:/ 38 /a/ 32 (70)	/a:/28 /a/ 47 (75)	/a/ 68 /a:/ 10 (78)	/a/ 69 /a:/ 8 (77)
/ʌ/	/ɒ/ 52 /ɔ/ 12	/ɒ/ 38 /ɔ/ 25	/o/ 55	/o/ 56
/o/	/ɒ/ 18 /ɔ/ 41	/ɒ/ 14 /ɔ/ 48	/o/ 67	/o/ 59
/u/	/ʊ/ 33 /u/ 29	/ʊ/ 37 /u/ 29	/ɯ/ 59	/ɯ/ 55
/ɯ/	/ʊ/ 22 /u/ 20 /ə/ 16	/ʊ/ 14 /u/ 17 /ə / 33	/ɯ/ 62	/ɯ/ 64

Table 4. Composite similarity index scores (numbers in the bracket represent combined scores)

With regard to K /ʌ/, the results of the identification experiment show roughly a chance level performance by most subject groups except for AE2. For K /ʌ/, AE1 has AE vowels /ɒ/ and /ɔ/ as a match with the similarity score of 52 and 12 respectively while AE2 has the score of 38 and 25 for the same AE vowels. Both J groups have J vowel /o/ as a match with the similarity score of 55 (J1) and 56 (J2). If a higher similarity index score directly translates to a better categorization (as was the case with K /a/), one would expect AE1, J1 and J2 to perform better than AE2 in the identification of K /ʌ/. However, the opposite is the case, where the best identification performance came from AE2. A possible explanation can be found by carefully examining the L1 categorization patterns represented in table 4 by each subject group and the articulatory description of K /ʌ/ and matched L1 vowels.

The L1 categorization pattern of AE and J groups seems to indicate that the 'backness' feature has more weight than the 'rounding' feature since the L1 categories chosen by all groups have [+ round] despite the fact that K /ʌ/ has [- round]. One may question this position by arguing that AE and J point vowel /a/ are no longer available as a possible match to K /ʌ/ since they are already well matched to K /a/. This may hold truth for J group. However, AE subjects could have chosen either AE /ɜ/ or /ə/, which they did not, as a match to K /ʌ/ if rounding would have more weight. Within the AE groups' categorization pattern for K /ʌ/, the similarity index scores given to AE /ɒ/ and /ɔ/ show a reverse trend where /ɒ/ gets a lower (from AE1's 52 to AE2's 38) and /ɔ/ gets a higher score (from AE1's 12 to AE2's 25). Between AE /ɒ/ and /ɔ/, in terms of the feature 'height' /ɔ/ would be a better match to K /ʌ/. The lower score given to AE /ɔ/ may be due to the fact that it is a better match to the another Korean vowel /o/ in terms of backness, rounding and height, hence not readily available as a match to K /ʌ/. The reverse trend in the similarity score given by AE groups, we believe, possibly entails two things. The higher score given to AE /ɔ/ may indicates that AE2 subjects are beginning to pay more attention to the height of K /ʌ/. More importantly, though, is the possibility that AE2 group may be aware of the lack of rounding in K /ʌ/. It is like saying 'this vowel is neither /ɒ/ nor /ɔ/ but something else'. In other words, AE2 group may be developing a more accurate characterization of K /ʌ/, and possibly signaling an onset of a new vowel category. If AE2 subjects are indeed developing a new vowel category (i.e. K /ʌ/), it would explain why the group has performed considerably better than the other groups in the identification and most of the discrimination tasks involving K /ʌ/.

With regard to K /o/, both AE groups gave a higher similarity rating to AE /ɔ/ than /ɒ/ indicating possibly their awareness that AE /ɔ/ is a better match in terms of vowel height. Given the identical binary specification between AE /ɔ/ and K /o/, one may wonder why the similarity score is not as high as the ones given to K /a/. A possible explanation would be the difference in vowel length. Vowel length is no longer contrastive in Korean and K /o/ would be more like AE /ɒ/ in length. The reverse trend in the similarity score between AE1 and AE2 (from 18 to 14 for /ɒ/ and 41 to 48 for /ɔ/) could indicate the awareness by AE2 subjects that vowel length is not crucial in Korean, thereby giving less weight to the length feature. This could again be seen as the evidence of developing a finer and more accurate characterization of an L2 vowel. A substantially better identification and discrimination score by AE2 group regarding K /o/ would provide some support for this view.

In the case of J group the most noticeable aspects, compared to the AE group, of the categorization pattern with regard to the Korean /ʌ/ and /o/ is that these two Korean vowels are matched to a single L1 category (i.e. J /o/). Given the identical distinctive feature specification between K /o/ and J /o/, one would expect a relatively high similarity index score for J /o/ as a match to K /o/ and a relatively low score for J /o/ as a match to K /ʌ/. This certainly seems to be the case with J1 group where J /o/ is matched to K /ʌ/ with the score of 55 while J /o/, to K /o/ with the score of 67. However, in the case of J2 group the similarity index scores given to J /o/ as a match to K /ʌ/ and /o/ are very close at 56 and 59. Given this L1 categorization pattern difference, one would expect a better identification performance from J1 than J2 with regard to K /ʌ/ and /o/. One would also expect a better discrimination score for K /ʌ/ - /o/ pairs from J1 group. However, the results from both experiments show that J2 group performed consistently better than J1 group. In fact the worst identification score from the data occurred with the J1 group's identification of K /o/ with the mean average percent correct score of 35.8%. Given the roughly a chance level performance by J2 group in the identification of K /ʌ/ and /o/, and an even worse performance by J1 group, it seems that J group as a whole is not attending well to the crucial rounding feature upon which K /ʌ/ and /o/ are contrasted.

Regarding K /u/ both AE groups put AE /u/ as a better match than AE /u/. This is predictable from the fact that AE /u/ is a central long vowel while K /u/ is a back vowel. The similarity score difference given to AE /u/ by AE1 and AE2 is only marginal. It is difficult to state that AE2 group has developed a more accurate characterization of K /u/ on that marginal difference and use it to explain as a basis for a substantially superior performance shown by AE2 group in the identification of K /u/. The low similarity index scores given to AE /ʊ/ and /u/ by both AE groups would indicate that K /u/ is not clearly categorized.

With K /ɯ/, there is a three way matching pattern (to /ʊ/, /u/ and /ə/) by AE groups but with very low similarity index scores. This could indicate that AE subjects are possibly treating K /ɯ/ as a vowel quite different from any AE vowels. One interesting point is the relatively high similarity index score given to /ə/ by AE2 group. This may indicate an emerging awareness by AE2 group that K /ɯ/ is unrounded. This could be an instance of hypercorrection where too much weight is given to rounding which overrides frontness and height since in terms of binary features K /ɯ/ would share more common features with AE /ʊ/ than /ə/. The fact that AE2 has the best identification score for K /ɯ/ may again indicate that AE2 subjects are in the process of forming a new vowel category, albeit incomplete. The lack of clear categorization with regard to K /u/ and /ɯ/ by both AE groups is clearly linked to the similar discrimination score level (89.7% and 88.8% correct) obtained for K /u/ - /ɯ/ pair.

In the J group's categorization of K /u/ and /ɯ/, we have a pattern which is similar to the J group's categorization of K /ʌ/ and /o/ where two L2 vowels are categorized with the same L1 vowel, in this case with J /ɯ/. In terms of binary feature specification, K /ɯ/ is identical to J /ɯ/. This is reflected in the higher similarity index scores given by both J groups to J /ɯ/ as a match to K /ɯ/ than K /u/. However, J2 groups has a larger similarity score difference than J1, possibly indicating J2 subjects are being more aware of the rounding contrast between K /u/ and /ɯ/. Since K /ɯ/ is a relatively well categorized vowel for both J groups, one would expect a similar level of discrimination performance in distinguishing K vowel pairs in which one is K /ɯ/. The discrimination results for such pairs indeed show similar level of performance by both J groups except for K /u/ - /ɯ/ in which J2 performed substantially better. Given the fact that J2 group possibly has better awareness of rounding contrast between K /u/ and /ɯ/, this would be expected. In the identification performance regarding K /ɯ/ both J groups has the same score. However, with K /u/ J2 group performed much better than J1 (78.3% and 48.3% correct) in identification, providing a support for the explanation put forward.

What emerges from the discussion above is the crucial nature of the perceptual mapping between the target L2 vowels and their matching L1 counterparts. One dimension of this perceptual mapping would be the link between an L2 and an L1 vowel and with what degree of similarity (e.g. identification of K /a/ and other four Korean vowels by all groups). But the more important dimension, as far as the present study is concerned, seems to be the match pattern. More specifically, the pattern can be one to many (e.g. K /ʌ/ to AE /ɒ/ and /ɔ/, K /ɯ/ to AE /ʊ/, /u/ and /ə/) and many to one (e.g. K /ʌ/ and /o/ to J /o/, K /u/ and /ɯ/ to J

/ɯ/). The results from the present study showed that the cross-sectional and cross-language difference along these two dimensions of perceptual mapping underpins the performance difference in identification and discrimination among the subject groups and explains the significance of L1-Vowel factor interaction and L1-Vowel Contrast factor interaction in the ANOVA results. The results also indicate that, although less obvious than its effect on identification, L1 does influence the discrimination performance (e.g. performance difference in K /u/ - /ɯ/ pair contrast by AE and J subject groups).

One interesting conjecture that can be made by observing the different perceptual matching patterns between AE and J groups is that a one to many pattern (i.e. AE group) could be a better starting ground for L2 vowel acquisition than a many to one pattern (i.e. J group). The cross-sectional difference in the identification of K /ʌ/ and /ɯ/ between AE and J group showed that there is a considerable improvement from AE1 to AE2 while there is a very small (with K /ʌ/) or no improvement at all (with K /ɯ/) from J1 to J2. There is a similar magnitude of cross-sectional improvement in the identification of K /u/ and K /o/ between AE and J group. However, AE2 has 90.3% and 79.2% correct identification for K /u/ and K /o/ while J2 has 78.3% and 51.7% for K /u/ and K /o/. It is not clear whether the many to one mapping pattern of J group actually put a ceiling effect in the learning of certain Korean vowels but at least it seems to be preventing J subjects from acquiring the Korean vowels at the same rate as the AE subjects. In a sense, having a more crowded L1 vowel space could be like having more angles to examine an entity, thereby one can quickly and more accurately characterize that entity.

To conclude, the results from the present study clearly showed that the nature of perceptual mapping between the target L2 vowels and their matching L1 counterparts is crucially related to the ways in which L2 sounds are accommodated by the learners with different L1 background. The study results provided evidence that L1 exerts effect on both the identification and discrimination in L2 vowel perception. It was also shown that different mapping patterns (i.e. one to many and many to one) not only affect the initial stage of L2 sound accommodation but may also have crucial implication for the developmental aspects of L2 learning.

Although not discussed in this paper, we are currently investigating how the perceptual mapping patterns in the data can be linked to the combined (i.e. AE, K and J) F2-F1 by F1 Bark scaled vowel space. The data in this study, especially the discrimination one, is also being used as a testing ground for Best's (1995) Perceptual Assimilation Model which was proposed to account for variability in discrimination of non-native sound contrasts.

REFERENCES

Best, C. T. 1995. A direct realist view of cross-language speech perception. in W. Strange (ed.), *Speech Perception and Linguistic Experience, Issues in Cross-Language Research.* Baltimore: York Press.

Polka, L. 1995. Linguistic influences in adult perception of non-native vowel contrasts. *Journal of the Acoustical Society of America,* 97-2, 1286-1296.

Stevens, K., Liberman, A., Studdert-Kennedy, M., & Ohman, 5. 1969. Cross-language study of vowel perception. *Language and Speech,* 12, 1-23.