한국어 3중 대립 음소에 대한 일본인의 지각적 범주화*

김지현 · 김정오
서울대학교 인지과학 연구소
kimjeehyun@cogsci.snu.ac.kr · jungokim@plaza.snu.ac.kr

Japanese Adults' Perceptual Categorization of Korean Three-way Distinction

Jeehyun Kim · Jung-Oh Kim
Institute of Cognitive Science, Seoul National University
kimjeehyun@cogsci.snu.ac.kr · jungokim@plaza.snu.ac.kr

Abstract

Current theories of cross-language speech perception claim that patterns of perceptual assimilation of non-native segments to native categories predict relative difficulties in learning to perceive (and produce) non-native phones. Perceptual assimilation patterns by Japanese listeners of the three-way voicing distinction in Korean syllable-initial obstruent consonants were assessed directly. According to Speech Learning Model (SLM) and Perceptual Assimilation Model (PAM), the resulting perceptual assimilation pattern predicts relative difficulty in discrimination between lenis and aspirated consonants, and relative ease in the discrimination of fortis. This study compared the effects of two different training conditions on Japanese adults'perceptual categorization of Korean three-way distinction. In one condition, participants were trained to discriminate lenis and aspirated consonants which were predicted to be problematic, whereas in another condition participants were trained with all three classes of consonants. Results indicated that the latter condition was more effective and that the relative "learnability" did not seem to depend lawfully on the perceived cross-language similarity of Korean and Japanese consonants.

* 본 연구는 한국학술진흥재단 기초학문 육성지원 사업(과제번호 KRF-2003-074-HS0003)의 지원을 받아 수행되었음.

interest in cross-language comparisons of phonetic categories, growing out of the well-documented problems that adult second language (L2) learners have in acquiring a new phonological system. In his Speech Learning Model (SLM), Flege (1995) claims that continuing problems with "accented" production of phonetic segments can be attributed in large part to L2 learners' representation of the L2 segments as equivalent to "similar"segments in the native language (L1). That is, if the L2 phones are sufficiently similar to L1 phones, they will be perceptually assimilated to those native categories, with the result that both L1 and L2 segments are produced differently from native monolingualspeakers'utterances. If, however, L2 phones are sufficiently dissimilar from any L1 categories (i.e., "new"), the L2 learner will (eventually) establish distinct L1 and L2 phonetic categories, and production of the L2 segments will become more native-like.

In her Perceptual Assimilation Model (PAM), Best (1994, 1995) also invokes the concept of cross-language phonetic similarity to predict the relative difficulties that listeners will have in perceptual differentiation of non-native segmental contrasts. She describes several patterns of perceptual assimilation of L2 segments to L1 phonological categories, which are determined by the perceived phonetic similarity of L1 and L2 segments. Two L2 segments which are judged as equally "good"instances of a single L1 category (Single-Category pattern) will be most difficult to differentiate, while two L2 segments that are assimilated to two different L1 categories (Two-Category pattern) will be very easy to discriminate. In addition, contrasting L2 segments

that differ in their judged goodness as instances of a single L1 category (Category-Goodness pattern) will yield intermediate levels of perceptual difficulty. Finally, if an L2 segment is sufficiently dissimilar from any L1 category, it may be considered an "uncategorizable"speech sound. When paired with another L2 phone that is phonetically similar enough to be categorized as an instance of an L1 category, the two phones will be easily discriminated.

Experiment 1

According to both models, the perceived similarity of segments in L1 and L2 is an important determiner of the pattern of initial perceptual problems and persistent learning difficulties adult L2 learners have in mastering the L2 phonological system. The purpose of Experiment 1 was to assess the perceived relation between Japanese and Korean syllable-initial obstruent consonants.

Method

Participants

Tennative Japanese speakers living in Korea participated. All were students of Language Education Institute in Seoul National University. No participant reported a history of hearing or speech disorders.

Speech materials

Thirteen male native speakers each of Korean and Japanese produced Korean or Japanese consonants followed by /a/ in a carrier phrase. The carrier phrases were "이것은 ____다," for Korean and "これは ____です," meaning "This is ____" for Japanese. The consonants used were Korean /p, p^h, p*, t, t^h, t*, k, k^h, k*, tç, $tç^h$, tç*, s, s*, h/ and Japanese /b, p, d, t, g, k, dz, s, d, tç, h/.

Recordings were made, in a sound booth, digitally with a Shure SM48 microphone through CoolEdit 2000 into a SONY-PSM R300. Sampling rate was set at 44,100 Hz. The distance between the microphone and the speaker was about 10 cm. Praat ver. 4.2 and Sound Forge ver.7.0 were used for editing the stimuli.

To prevent vowel length from being used as a cue in consonant discrimination, all vowels were truncated to the same duration. To minimize a "clipped" percept at the vowel ending, the intensity of the last 20 ms of the vowel ending was ramped off from 100% to 0%. All stimuli were set to 65dB.

To ensure that only good examples of the Korean and Japanese consonants were used as stimuli, four native Korean speakers from Seoul and four native Japanese speakers who had been in Korea for less than six month, all from the Tokyo area, judged consonants from their native language. The best five (out of thirteen) tokens, as judged by native speakers, were selected for use as stimuli.

Procedure

Participants were first asked to identify each token as an instance of some Japanese consonant category, and then to rate the token for goodness-of-fit to the (just-selected) Japanese category using a scale ranging from bad example (1) to very good example (7). The choices of Japanese *Katakana* orthography (IPA representations given) were as follows:

カ/ka/, ガ/ga/, サ/sa/, ザ/dza/, タ/ta/, ダ/da/, チャ/ta/, ジャ/da/, ハ/ha/, バ/ba/, パ/pa/.

The participants were tested individually in a sound booth and heard the stimuli over Sennheiser HD212Pro headphones from a PC equipped with Sound Blaster Audigy 2 ZS Platinum Pro soundcard. They were instructed to identify and rate each token by clicking one of the eleven *Katakana* buttons and one of the seven scale buttons on the monitor.

Results and discussion

Table 1 represents the results for the Korean consonants. Unlike the Japanese consonants, there was not a "correct"classification for the Korean stimuli. Of the fifteen Korean consonants examined, eleven were consistently (>75%) classified as instances of a single Japanese consonant category. These consonants received a mean goodness rating of 3.8 for the corresponding Japanese category.

Two Korean consonants were constantly identified in terms of one Japanese consonant, and more interestingly, the goodness ratings for them were similar. Both Korean lenis and aspirated consonants were heard as Japanese voiceless consonants with a high level of goodness rating (3.6). (Japanese voiceless consonants got a mean goodness rating of 4.3.) Some Korean consonants, on the other hand, were identified in terms of two Japanese consonants. Participants heard these consonants as intermediate between two Japanese categories. Korean fortis washeard as Japanese voiceless consonants about 65% of the time, and as voiced consonants about 33% of the time.

According to Speech Learning Model (SLM) and Perceptual Assimilation Model (PAM), the resulting perceptual assimilation pattern should be as follows: Lenis and aspirated consonants showing Single-Category pattern will be most difficult to discriminate, while the "new" fortis consonants will be relatively easy for Japanese learners.

Consonant Stimuli	*Percent Identification and Rating*										
	パ pa	バ ba	タ ta	ダ da	カ ka	ガ ga	サ sa	ザ dza	チャ ta	ヅャ dZa	ハ ha
/p/	**97**	2							1		
	(4.0)	(1.5)							(4.0)		
/p^h/	**94**		2		1						2
	(4.9)		(1.0)		(1.0)						(4.0)
/p*/	**65**	**33**	1								1
	(3.3)	**(2.3)**	(4.0)								(5.0)
/t/	1		**96**	3							
	(4.0)		**(4.2)**	(3.3)							
/t^h/			**99**								1
			(4.6)								(4.0)
/t*/			**66**	**33**					1		
			(3.0)	**(2.1)**					(4.0)		
/k/					**98**	2					
					(3.7)	(1.0)					
/k^h/	1		1		**98**						
	(3.0)		(3.0)		**(3.2)**						
/k*/			1		**66**	**33**					
			(4.0)		**(2.3)**	**(1.5)**					
/s^h/							**100**				
							(4.7)				
/s*/				1			**96**	3			
				(1.0)			**(3.1)**	(1.5)			
/tC/									**88**	12	
									(4.2)	(2.0)	
/tC^h/									**99**	1	
									(4.3)	(3.0)	
/tC*/								6	**52**	**42**	
								(1.6)	**(2.8)**	**(2.5)**	
/h/	2	1	1				1				**95**
	(1.5)	(5.0)	(1.0)				(4.0)				**(4.6)**

Table 1. Mean percent identification and goodness rating (in parentheses) of Korean consonant stimuli in terms of Japanese categories.

Experiment 2

The purpose of Experiment 2 was to compare the effects of two different training conditions on Japanese adults'perceptual categorization of Korean three-way distinction.

Method

Participants

Twentynative Japanese speakers living in Seoul participated. No participant reported a history of hearing or speech disorders.

Speech materials

Five male native speakers of Korean produced nine consonant-vowel syllables assingle utterances. These were created by combining the three places of stop articulation (bilabial, alveolar, and velar) with the three voicing classes (lenis, aspirated, and fortis). Recording was made as in Experiment 1. Praat ver. 4.2 was used for editing, and allstimuli were set to 65dB.

Procedure

Participants tookpart in three sessions. The first and last sessions consisted of the pretest and posttest phases each with one training block, whereas the middle session consisted entirely of training. Pretest and posttest consisted of a difference rating task and a phonetic identification task. Training used a categorical discrimination test with feedback. During training, participants heard three stimuli per trial and were asked to pick the odd item. Thetest differed from traditional oddity tasks in that it incorporated "catch" trials consisting of three physically different tokens of the same consonant. This encouraged the participants to respond only to phonetically relevant differences, not to any auditorily detectable differences.

Participants were randomly assigned to two training conditions. In one condition, participants were trained to discriminate lenis and aspirated consonants which were predicted to be problematic from the two models of cross-language perception, whereas in another condition participants were trained with all three classes of consonants.

Results and discussion

Participants in each condition were divided into two groups according to their performances on the pretest. Only the low-level participants' identification scores are reported here (Figures 1 and 2). In the lenis-aspirated training condition, the average improvement in identification accuracy was only 3.5% from 75.6%(SD=7.4) on the pretest to 79.1%(SD=6.0) on the posttest. In the lenis-aspirated-fortis condition, however, the improvement was 15.8% from 68.6%(SD=6.8) to 84.4%(SD=5.5). There were significant effects of test ($F(1, 27)=11.842$, $p<.05$) and of phoneme categories($F(2, 27)=4.115$, $p<.05$).

Contrary to the predictions from the two models reviewed above, perception of fortis consonants was most difficult, and perception of aspirated consonants improved the most.

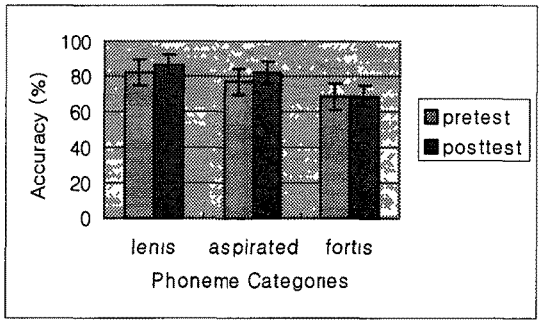

Figure 1. Identification scores in lenis-aspirated training condition.

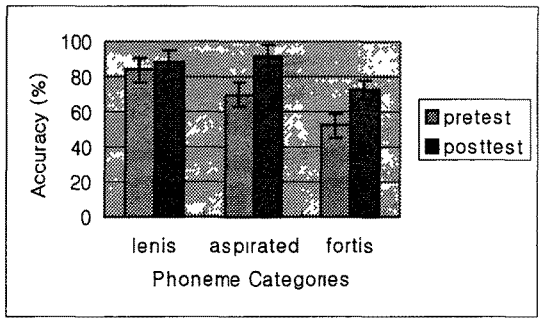

Figure 2. Identification scores in len is-aspirate -fortis training condition.

General Discussion

In lenis-aspirated training condition, the high-level participants' identification accuracy level for lenis consonants dropped in the posttest despite the fact that there had been very fewerrors during the training phases. This result indicates that the introduction of a new class, fortis at the posttest, can cause confusion for the L2 learners. It suggests that the traditional way of teaching phonemes using minimal pairs may not be the best one for Korean three-way phoneme distinction.It is possible that the participants learned to discriminate lenis and aspirated by attending to the difference in the amount of aspiration. Few errors at the last training block indicate that they almost mastered the distinction. The introduction of fortis which doesn't have aspiration could have caused the confusion between lenis and fortis. In the lenis-aspirated-fortis training condition, on the other hand, it is possible that fortis worked as an anchor with no aspiration so

that the participants can learn to attend both to the aspirated and tense properties of the three classes of sounds.

One important question is whether the relationship between cross-language mapping patterns in an unknown foreign language will also apply to individuals who are learning a second language. The results obtained here suggest that the PAM and SLM can not be readily extended to early stages of naturalistic L2 speech learning without further investigation. The degree of perceptual difference did not seem to depend on the extent to which the two members of a consonant contrast would be identified as instances of a single Japanese consonant category. The relative "learnability" of the non-native phonemes did not seem to depend lawfully on the perceived cross-language similarity of Korean and Japanese consonants.

Finally, this study suggests that L2 learners perceive non-native speech sounds on the basis of a new, though imperfect, perceptual criterions different from the ones they use for their native sounds. Thus, the induction of a missing dimension without considering the entire structure is less effective. Aproblem faced in non-native perceptual categorization should be approached from the entire perceptual structure established by their native language.

References

Best, C.T. 1994. The emergence of native-language phonological influence in infants: A perceptual assimilation model. In J.C. Goodman & H.C. Nusbaum, eds. *The Development of Speech Perception: The Transition from Speech Sound to Spoken Words*. Cambridge: MIT Press. 167-224.

Best, C.T. 1995. A direct realist view of cross-language speech perception. In W. Strange, ed. *Speech Perception and Linguistic Experience: Issue in Cross-Language Reseach*. Baltimore: New York Press. 171-204.

Flege, J.E. 1995b. Second language speech learning: Theory, findings and problems. In W. Strange, ed. *Speech Perception and Linguistic Experience: Issue in Cross-Language Research*. Baltimore: New York Press. 233-277.

Guion, S.G., J.E. Flege, R. Akahane-Yamada, & J.C. Pruitt. 2000. An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *Journal of Acoustical Society of America*107, 5, 2711-2724.