PERCEPTION OF JAPANESE PLOSIVES BY KOREAN SPEAKERS

Kanae AMINO,[1] Seong-Rim JI,[2] Shigeko SHINOHARA[3]

Dept. Electric. & Electr. Eng. Sophia Univ.,[1] Phonetics Laboratory, Sophia Univ. [2,3]
amino-k@sophia.ac.jp,[1] slj4152@hotmail.com,[2] shinohara@free.fr[3]

ABSTRACT

This paper focuses on the perception of Japanese voiceless plosives by Korean listeners to: 1) assess whether the categorization of utterance initial- and medial- Japanese plosives into Korean three-way laryngeal categories is the same as in loanwords from Japanese; 2) identify the acoustic cues relevant to the identified patterns. We used modified stimuli from nonsense word tokens of native Japanese speakers, manipulating F0, amplitude and temporal characteristics. Loanwords represented only a part of the perception patterns: utterance initial [p] was mostly heard as aspirated and medial [p] was identified as either aspirated or fortis. The acoustic cues of F0 on the following vowel and temporal characteristics particular to the positions had effects on laryngeal perception.

Keywords: Loanword, Korean, Japanese, plosive.

1. ISSUES

Korean has a well-known three-way laryngeal contrast while Japanese plosives have only [+/– voice] categories. In Korean loanwords, Japanese voiceless plosives are adapted in distinct phonological categories depending on whether they appear in word- initial or medial positions. Another intriguing fact is that the adaptations of medial plosives differ depending on the source. The aim of this paper is to use experimental data to study the perception of Japanese voiceless plosives by Korean speakers.

1.1. Laryngeal categories and acoustic cues

Of the many analyses proposed for Korean three-way laryngeal contrast (see [7]), we adopted the distinction: lenis [-tense, -spread glottis (s.g.)], fortis [+tense, -s.g.], and aspirated [+tense, +s.g.] by [6], as it is relevant for the loanword patterns discussed in 1.3. The laryngeal contrasts in Korean are realized with multiple phonetic cues: pitch values on the following vowel signal [+/-tense] distinction [5]; VOT and closure duration are involved in distinguishing within the [+tense] categories; the low spectral tilt signals fortis consonant [5]. A further complication is that temporal characteristics such as VOT and closure duration values vary considerably depending on the type of consonants and their position in a phrase [1][2]. Another relevant fact for the categorization of voicing contrast as a whole is that lenis is realized as voiced only in the post sonorant position in Korean. In contrast, Japanese is a moraic language where durational variance is relatively stable [8], but F0 value has a distinctive role for the lexical pitch accent. Therefore, there may be differences in interpretation of phonetic characteristics between the two languages.

Accordingly, loanword adaptation patterns of Japanese plosives into Korean differ from those of English and French plosives. In addition Japanese loanwords also show an interesting positional asymmetry.

1.2. Positional asymmetry in Japanese loans

The original [+/-voice] is neutralized in the initial position in loanwords and distinguished in the word medial position. The word initial Japanese voiceless and voiced plosives, and word medial voiced plosives are uniformly adapted as lenis plosives across available data [3][4].

(1) Korean adaptation of Japanese voicing contrast

(Data from [4] based on common usage. |Japanese input|; /Korean output/. (’) indicates fortis plosives.)

Jp initial voiceless plosive to Kr lenis

|**t**ojota| → /**t**ojot’a/ ~ /**t**ojot^{h}a/ ‘Toyota’
|**k**arate| → /**k**alat’e/ ~ /**k**alat^{h}e/ ‘Japanese martial art’
*There is no single /p/ in native Japanese words.

Initial or medial voiced plosive to lenis

|**b**uçi| → /**p**usi/ ‘soldier’
|**d**açi| → /**t**asi/ ‘soup stock’
|so**b**a| → /so**p**a/ ‘Japanese buckwheat noodle’
|na**g**oja| → /na**k**oja/ ‘Nagoya’

Word medial plosives make opposition in loans. However, there is inconsistency in the adaptation data: the following patterns have been reported for adapting medial voiceless plosives to fortis and aspirated [4]:

(2) Medial voiceless plosive to fortis or aspirated

|nari**t**a| → /nali**t'**a/~ /nali**t^h**a/ 'Narita (Airport)'
|oosa**k**a| → /osa**k'**a/~ /osa**k^h**a/ 'Osaka'

Another study reports lenis rendition for medial |t| and fortis rendition for |k| [3]. These data were based on the Korean Etymological Dictionary and assumed to be the actual pronunciation. Given that loanword adaptation might involve factors other than perception of acoustic properties, our first task is to verify whether the perception is reflected in the Korean loanword adaptation of Japanese.

1.3. Distinct categories for English and French

The second issue is to assess the reasons for the initial neutralization of Japanese voicing contrast. An interesting puzzle arises when we observe Korean adaptations of English and French words, which maintain the original voice distinction in every position. Below is data drawn from [6].

(3) Korean adaptation of French voicing contrast

|**p**aRi| → /**p^h**ali ~ **p'**ali/ 'Paris'
|na**p**ɔleɔ̃| → /na**p^h**oleoŋ ~ na**p'**oleoŋ/ 'Napoleon'
|**b**ɔR**d**o| → /**p**olɯ**t**o/ 'Bordeaux'

Here French voiceless plosives are rendered as both fortis and aspirated. On the other hand, adaptations of English words show an alternation between lenis and fortis (the latter in an emphatic expression) of voiced plosives; while voiceless plosives are adapted to aspirated in every position.

(4) English into Korean

|**p**aɪp| → /**p^h**ai**p^h**ɯ/ *pipe*
|s**p**i:**k**ə| → /sɯ**p^h**i**k^h**ə/ *speaker*
|**b**ɑ:| → /**p**a ~ **p'**a/ *bar* (the latter with emphasis)

These patterns suggested feature mapping during the adaptation process: [+/-tense] to [+/-voice] in French, and [+/-s.g.] to [+/-voice] in English [6].

[3] reported no variation in such cases: French voiceless stops are adapted as Korean fortis; English voiced stops as lenis. Therefore, the question is why Japanese initial voiceless plosives are adapted as lenis while French ones are adapted as fortis, when both languages have only lightly aspirated voiceless plosives.

Korean loanword data from different languages indicate that multiple factors are involved in identification of laryngeal contrast of L2. We investigated the perception of Japanese plosives by Korean listeners by controlling F0, amplitude and temporal characteristics. The aim was to: 1) assess how Korean listeners categorize utterance initial- and medial- Japanese plosives into Korean three-way laryngeal categories; and 2) identify the acoustic cues relevant to the patterns found. In particular, we are interested in Korean listeners' sensitivity to positional difference in Japanese initial and medial plosives.

2. EXPERIMENT

The experiments consisted of acoustic analyses of original Japanese sequences and perception tests of the natural and modified stimuli.

2.1. Experimental design

2.1.1. Original stimuli

Five male speakers of Tokyo Japanese uttered the two-syllable nonsense phrase [papa] five times with HL tone pattern (due to the initial default accentuation). The test phrase was uttered in isolation in order to be closer to natural settings. Since bare nouns and verbs can often occur in the phrase initial position in both Japanese and Korean, Korean speakers may often hear and use Japanese lexical words in this position. The utterances were recorded in a sound-proof room onto digital audiotape and digitized at 16 kHz with 16 bit. Three out of five tokens were retained for the listening tests.

VOT, closure duration and F0 values of the relevant segments in the first syllable (S1) and second syllable (S2) were measured on waveforms and spectrograms on Praat 4.4.07. Average values and closure duration ratio are shown in Table 1.

Table 1: Acoustic characteristics of Japanese [papa].

	S1	S2
Mean VOT (Sd.)	32.7 ms (6)	14.2 ms (6)
F0 values of vowel (Sd.)	153.33Hz (25.23)	90.36Hz (19.52)
Mean Closure duration (Sd.)	n.a.	63.4 ms (15)
Closure duration ratio (Sd.)	n.a.	18.6 % (3.7)

2.1.2. Manipulated stimuli

We created 10 types of stimuli to discern the influence of each of the following parameters: 1) temporal characteristics in S1 and S2; 2) F0 patterns; 3) amplitude difference between S1 and S2. Only the 7 types shown in Table 2 were introduced in this study. The stimuli were created by the methods described in Table 3.

Table 2: Stimulus Variation (7 conditions)

ID	Condition	F0 contour	Amplitude
Set A: Original order (S1S2)			
#1	Original	HL	A1>A2
#2	Normalised amplitude	HL	A1=A2
#3	Nomalised F0	MM	A1>A2
#4	Reversed F0	LH	A1>A2
#5	Normalised amp & F0	MM	A1=A2
Set B: Reversed order (S2S1)			
#6	Original	LH	A1<A2
#7	Normalised amp. & F0	MM	A1=A2

Table 3: Stimulus Creation

ID	Methods
#1	Originally recorded sound files
#2	Original stimuli were modified using rms amplitude data to give both syllables the same average value.
#3	Original stimuli were modified using Praat F0 manipulation to give both syllables the same average value.
#4	Original stimuli were modified using Praat F0 manipulation so that the F0 values of the first and second syllables were transposed.
#5	Manipulation #3 was applied to stimuli #2.
#6	Created by transposing S1 and S2 of stimuli #1.
#7	Created by transposing S1 and S2 of stimuli #5.

In order to create stimuli #2, amplitude (rms) was calculated for each syllable by using the computer program MATLAB. Then, coefficients *m* and *n* were calculated to give an average for both amS1 and amS2, Aavr. Thus,

(5) $n*A1 = m*A2 = Aavr$.
n, *m*: coefficients
A1: RMS amplitude of the first syllable
A2: RMS amplitude of the second syllable
Aavr: average of A1 and A2

Finally, each element in S1 and S2 were multiplied by the coefficients *n* and *m*, respectively.

For stimuli #6 and #7 transposition between the two syllables was conducted as follows: first, the intervals (the start and the end) for S1 and S2 were determined based on the waveforms, and the duration of S2 closure was measured. The S2 interval was extracted and put before S1; a silence was then added between S2 and S1 that had the same duration as the closure of S2.

2.1.3. Listeners and procedure

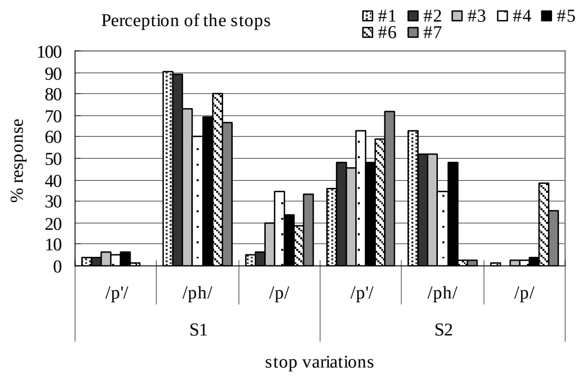
Five native Korean speakers with normal hearing participated in the perception test. The participants were university students residing in Tokyo for an average of 7.6 years and had been learning Japanese for an average of 5.6 years.

Participants listened to the stimuli diotically through headphones (SONY MDR-CD900ST). The stimuli were originally created for three consonants [p, t, k], although here we only report the results relating to the bilabial plosive. The experiment was divided into two parts: each part had 450 stimuli (corresponding to 3 tokens for 5 speakers, 10 conditions, and 3 consonants) presented in a pseudo-random order. The participants categorized the first consonant in part 1 and the second consonant in part 2. They were not told the purpose of the experiment; consequently they were mostly unaware of the origin of the stimuli.

2.2. Results

The results of the perception tests are shown in Figure 1.

Figure 1: Percentages of responses for each condition. (Original syllable order is retained in conditions #1~#5, and in #6 and #7 it is reversed.)



The data were submitted to statistical analyses.

In order to observe differences between the three stops in the original stimuli, the Kruskal-Wallis test was performed. The results show that differences were significant in both S1 and S2 ($p < 0.001$). Further comparison using a Mann-Whitney *U*-test showed that the number of Ph responses was greater than P’ and P in S1. In S2, the order of the responses was Ph>P’>P; all differences were significant.

A one-way ANOVA was performed to see the effects of F0 on the perception of the stops (differences between the conditions in #1, #3, #4 and #5, where F0 was modified). Significant differences in Ph and P were observed for S1, and in Ph for S2 ($p < 0.001$). P’ of S2 also showed significant differences ($p < 0.05$). Tukey’s post-hoc test showed significant differences between the following pairs: {#1:#4} and {#1:#5} for Ph and P in S1, and for Ph in S2; {#1:#4} for P' in S2.

In the stimuli with exchanged syllables in normalised conditions {#5:#7} the Mann-Whitney *U*-test showed significant differences in P for S1 ($p < 0.05$) and in P', Ph and P for S2 ($p < 0.01$).

No effects of amplitude (rms) on the perception of the stops were observed. Paired *t*-tests showed no significant differences in S1 or S2 between {#1:#2} or {#3:#5}. However, there was a slight decline in P' and Ph in #2.

We can draw the following conclusions from the statistical analyses. First, across all the stimulus types, the majority of S1 (the utterance initial Japanese |p|) were perceived as /p^h/, whereas the perception of S2 changed depending on the position where it appeared. In particular, the perception shifted to /p'/ and /p/ from /p^h/ in the reversed stimuli. Second, by lowering F0 the perception of /p^h/ decreased and that of /p/ increased in the initial syllable. As for the second syllable, there was a positive effect of F0 on perception of /p'/ ({#1:#4}), but not for /p^h/. Third, as expected by negative evidence, the amplitude difference due to the pitch accent had little role in Korean plosive categorization.

3. DISCUSSION

Categorization of Japanese utterance initial |p| as /p^h/ in our perception test differed from the loanword data where it was adapted as lenis. The medial |p| was categorized either as aspirated /p^h/ or fortis /p'/, the same as in one of the loanword data [4], but not for the data based on the Etymological Dictionary [3]. Thus, loanword adaptation does not agree entirely with perception patterns of Korean laryngeal categorization; phonological adaptation seems to involve different processes.

There were clear differences in ratios of temporal characteristics depending on position, based on the results of the cross-spliced manipulated stimuli. The majority perception of S1 as /p^h/ in the original order seems to be related to the longer VOT in S1 in the original stimuli. When S2 was transposed to the initial position the reason for perception as /p'/ rather than as /p^h/ seems also to be related to VOT differences between S1 and S2 in the original stimuli. The VOT in S1 (33 ms) is about 2.4 times greater than that in S2 (14 ms). A Korean study showed that the average VOT of utterance initial /t^h/ was 58 ms whereas medial VOT ("syllable initial level") was only 15 ms; it also reported the VOT of utterance initial /t/ as 30ms [2:173]. In a perception study in Korean, cross-spliced /p/ embedded in "syllable level" was misperceived as /p^h/ [1]. In our study, it seems that 14 ms was not long enough even for a lenis in the initial position, and resulted in an increase of /p'/ categorization.

F0 lowering incurred a shift from aspirated toward lenis percept. It is often reported in the literature that higher F0 is a cue for [+tense] in Korean [5]. Thus, in our study the changes in perception between /p/ and /p^h/ of S1 in F0 manipulated stimulus pairs could be predicted. We raise a question on the impact of pitch accent in Japanese. Japanese words can have HL, LH, or LHL tone patterns. In the loanword data this difference seems to be ignored, implying the listeners' capacity to deduce this information out of consonant identity calculation. The reason why S1 with higher F0 was not identified often as /p'/ must be due to its longer VOT. Reasons for the slight shift in identification of S2 /p^h/ toward /p'/ due to F0 rise is not clear.

In sum, this study suggests that only a part of perceived information is selected for phonological adaptation.

4. REFERENCES

[1] Baker, K. 2002. Prosodic structure and perception of Korean domain-initial coronal stops, *International J. Korean Linguistics* 11, 119-132.

[2] Cho, T., Keating, P. 2001. Articulatory and acoustic studies on domain-initial strengthening in Korean, *J. Phonetics* 29, 155-190.

[3] Ito, C., Kang, Y., Kenstowicz, M. 2006. The Adaptation of Japanese loanwords into Korean. *MIT Working Papers in Linguistics* 52, 65–104.

[4] Kim, H-S., Shinohara, S. 2006. A feature-driven non-native percept in the loanword adaptation between Japanese and Korean, *Japanese Phonology Forum 2006 Aug. 23-25 Waseda University, Tokyo*. Handout.

[5] Kim, M-R. Bedder, P.S., Horrocks, J. 2002. The contribution of consonantal and vocalic information to the perception of Korean initial stops. *J. Phonetics* 30, 77-100.

[6] Kim. H-S. 2005. The representation of the three-way laryngeal contrast in Korean consonants. In van Oostendorp, M., van de Weijer, J. (eds.), *The internal organization of phonological segments*. Berlin New York: Mouton de Gruyer, 287-315.

[7] Kim., M-R., Duanmu, S. 2004. "Tense" and "lax stops in Korean, *J. East Asian Linguistics* 13, 59-104.

[8] Takeda, K. Sagisaka, Y. Kuwabara, H. 1989. On sentence-level factors governing segmental duration in Japanese, *J. Acoust. Soc. Am.* 86, 2081-2087.