PERCEPTION OF INTERVOCALIC CONSONANT CLUSTERS BY JAPANESE LISTENERS

Masaki Noguchi [1]
[1]University of British Columbia

1. Introduction

Perceptual epenthesis by Japanese listeners has been known as a quintessential example of phonological bias in speech perception. Since Japanese has a phonotactic constraint that prohibits the occurrence of non-geminate consonant clusters, it has been argued that Japanese listeners upon hearing such consonant clusters tend to perceive an illusory vowel /u/ between the consonants so that their percepts conform to the phonotactic constraint. Experimental studies have demonstrated that because of this perceptual epenthesis Japanese listeners have significant difficulty in perceiving the difference between stimuli that contain a non-geminate consonant cluster (e.g. ebzo) and their counterparts with an intervening /u/ (e.g. ebuzo) [1].

The prohibition of non-geminate consonant clusters, however, is not absolute in Japanese. Some clusters are actually found in the surface phonetic form of naturally produced utterances as a result of vowel devoicing, which is usually described as the high vowels /i, u/ being devoiced when they occur between voiceless consonants. However, the actual outcome of the process ranges from devoicing to deletion [2], and its occurrence is rather probabilistically determined by various phonological factors, such as vowel type (e.g., high vowels are more likely to be devoiced than non-high vowels) and consonant type (e.g., devoicing is more likely to occur next to voiceless consonants than voiced consonants) [3].

Since vowel devoicing is an allophonic rule, it is assumed that once Japanese listeners learn it they should be able to recover underlying vowels whenever they hear consonant clusters that may have arisen through vowel devoicing. However, studies have demonstrated that it is not the case. Recovery does not happen unless it is lexically supported. In other words, Japanese listeners seem to retain the phonetic representation of surface consonant clusters if their lexical knowledge does not motivate the recovery of underlying vowels [3]. This is further supported by a finding that Japanese listeners actually have access to the phonetic representation of surface consonant clusters in a speech perception task. When they were asked to rate the goodness of non-word stimuli with surface consonant clusters (phonotactically non-canonical but phonetically natural stimuli) and their non-word counterparts with an underlying vowel (phonotactically canonical but phonetically unnatural stimuli), they rate the former as better than the latter [4]. This suggests that Japanese listeners have phonetic knowledge about surface consonant clusters.

The goal of this study is to understand how Japanese listeners' phonetic knowledge about surface consonant clusters relates to perceptual epenthesis. It examines whether the perceptibility of non-geminate consonant clusters for Japanese listeners is affected not only by their phonotactic knowledge about the prohibition of those clusters, but also by their phonetic knowledge about how likely those clusters are to occur as a result of vowel devoicing. In order to do that, 19 Japanese listeners were tested on discrimination between non-geminate consonant clusters and their counterparts with an intervening /u/. Crucially, 4 different types of consonant clusters with different degrees of likelihood of occurrence as a result of vowel devoicing were tested. If perceptual epenthesis is solely triggered by phonotactic knowledge, there would be no effect of cluster types.

2. Methods

2.1. Design

An AX discrimination paradigm was used. In experimental trials, participants compared non-word stimuli with intervocalic non-geminate consonant clusters (VCCV) and their counterparts with an intervening /u/ (VCuCV). Consonant clusters were classified into 4 types according to the likelihood of their occurrence as a result of vowel devoicing; from lowest likelihood to highest, these were voiced stop - voiced stop (DD), voiceless stop - voiced stop (TD), voiceless stop - voiceless stop (TT), and voiceless fricative - voiceless stop (ST) [3] (Table 1). In filler trials, participants compared stimuli with a mid short /u/ (VCuCV) and their counterparts with a mid long /u:/ (VCu:CV).

Cluster type	Stimuli	Likelihood
DD	igdo, egba, ibgo, ebda	Low
TD	ikdo, ekba, ipgo, epda	Mid-low
TT	ikto, ekpa, ipko, epta	Mid-high
ST	isto, espa, isko, esta	High

Table 1: Cluster types.

2.2. Stimuli

A trained phonetician whose native language is Japanese produced stimuli with a mid short /u/ (e.g., igudo). Stimuli were normalized in vowel duration and intervocalic interval duration. The mid vowel was spliced out to generate stimuli

* msk_ngch@yahoo.com

with consonant clusters (e.g. igdo) or lengthened to generate stimuli with a long /u:/ (e.g. igu:do).

2.3. Procedure

In each trial, participants heard a pair of stimuli and decided whether the stimuli were two different things or two repetitions of the same thing. The ISI was set to 500 msec. A block consisted of 64 experimental trials (32 same trials and 32 different trials) and 64 filler trials (32 same trials and 32 different trials). The block was repeated twice with an intervening short break.

3. Results and discussion

Performance was measured in terms of sensitivity (d prime) separately for filler trials (Figure 1) and experimental trials (Figure 2).
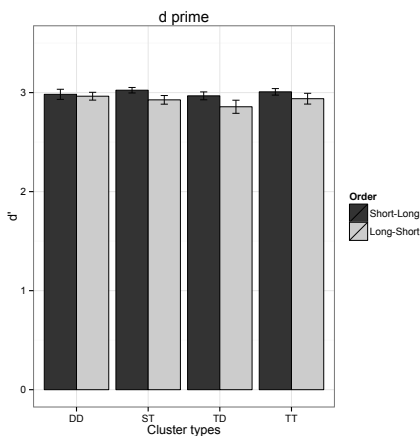

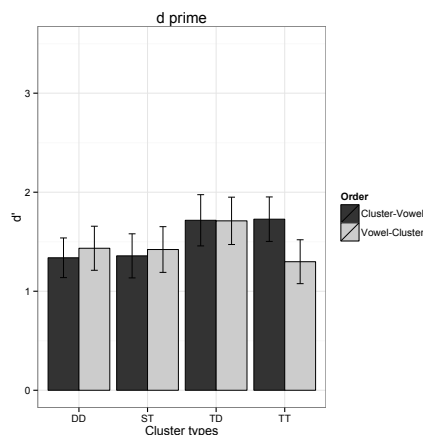**Figure 1**: d prime scores for filler pairs

Figure 2: d prime scores for experimental pairs.

Cluster type and order of stimuli presentation were taken into analyses. D prime scores for the filler trials show that Japanese listeners have good sensitivity to the vowel duration contrast. A repeated measures ANOVA revealed no main effect or interaction. D prime scores for the experimental trials show that Japanese listeners have poor sensitivity to the contrast between consonant clusters and their counterparts with an intervening /u/. A repeated

measures ANOVA revealed a main effect of cluster type $(F(3, 54)=2.77, p<0.05)$; d prime scores were numerically lower for DD and ST cluster types than TD and ST cluster types. The difference, however, did not follow the order of likelihood of occurrence. Interestingly, Japanese listeners had more difficulty in perceiving clusters of the least likely type (DD) and the most likely type (ST). This can be potentially explained by biases coming from two different kinds of phonological knowledge, one about the phonotactic constraint and the other about the allophonic rule. Clusters of the DD type were hard to perceive due to perceptual epenthesis, and clusters of the ST type were hard to perceive due to the recovery of the underlying /u/. Clusters of the TD and TT types were relatively easier to perceive. This is probably because these types are relatively immune to the phonological biases. The TD and TT types are more frequent than the DD type and do not trigger perceptual epenthesis as strongly as the DD type does, while the TD and TT types are less frequent than the ST type and do not trigger the recovery of underlying /u/ as strongly as the ST type does.

The ANOVA also revealed a marginal interaction between cluster type and order of stimuli presentation $(F(3, 54)=2.497, p=0.0694)$. A post-hoc analysis revealed that for the TT type, d prime scores were significantly higher when stimuli were presented in the order of VCCV and VCuCV. Previous studies have argued that in an AX paradigm discrimination is relatively easier when stimuli are presented in the order of unfamiliar and familiar items [5]. If the ordering effect is interpreted in terms of stimulus familiarity, what defines familiarity in this context is phonotactic canonicality.

4. Conclusion

The results showed that the perceptibility of non-geminate consonant clusters varied depending on cluster type, suggesting that the phonotactic constraint is not the only factor that affects the perception of non-geminate consonant clusters. A further study is needed to understand the mechanism behind the processing of non-geminate consonant clusters by Japanese listeners.

References

[1] Dupoux, E., Kakehi, K., Hirose, Y., Pallier, C., & Mehler, J. (1999). Epenthetic vowels in Japanese: A perceptual illusion? Journal of Experimental Psychology: Human Perception and Performance, 25(6), 1568-1578.

[2] Ogasawara, N., & Warner, N. (2009). Processing missing vowels: Allophonic processing in Japanese. Language and Cognitive Processes, 24(3), 376-411.

[3] Maekawa, K., & Kikuchi, H. (2005). Corpus-based analysis of vowel devoicing in spontaneous Japanese: an interim report. In J. van de Weijer, K. Nanjo, & T. Nishihara, eds., Voicing in Japanese, Mouton de Gruyter, pp 205-228.

[4] Fais, L., Kajikawa, S., Werker, J., & Amano, S. (2005). Japanese listeners' perceptions of phonotactic violations. Language and Speech, 48(2), 185-201.

[5] Tsushima, T., Shiraki, S., Yoshida, K., & Sasaki, M. (2003). On stimulus order effects in discrimination of nonnative consonant contrasts. Acoustical Science and Technology, 24(6), 410-412.