THE EFFECT OF LEXICAL COMPETITION ON REALIZATION OF PHONETIC CONTRASTS: A CORPUS STUDY OF THE VOICING CONTRAST IN JAPANESE

Mafuyu Kitahara[1], Keiichi Tajima[2], and Kiyoko Yoneyama[3]

[1]Sophia University, [2]Hosei University, [3]Daito Bunka University
mafuyu@sophia.ac.jp, tajima@hosei.ac.jp, yoneyama@ic.daito.ac.jp

ABSTRACT

Studies on phonetic realization of phonological structure have revealed that certain lexical properties are reflected in fine details of speech production. Examples are vowel space expansion for words with dense neighbors, and VOT increase in voiceless stops for words with voicing minimal pairs. These lexical effects are primarily found in English, but there is much to explore in phonemically and prosodically different languages, such as Japanese. The present study investigated the duration of intervals around the burst of word initial velar stops in the Corpus of Spontaneous Japanese which has about 200 speakers' annotations given by trained phoneticians. Minimal pair competitors and word familiarity data are drawn from an 80,000-word database. Results show significant effects of competitors on the duration of the following vowel and word familiarity on the closure duration, which suggest that lexical effects do exist in Japanese, but in a radically different manner from English.

Keywords: speech corpus, voicing contrast, lexical competition, word familiarity, Japanese.

1. INTRODUCTION

Phonetic realization of phonological contrasts is influenced by lexical effects, such as word frequency/familiarity and neighborhood structure. Neighborhood is approximately calculated by the words that are in the range of one-phoneme difference by addition, deletion or substitution [4], and the number of words in a particular neighbourhood is called neighborhood density. Lexical effects are well attested in English, not only in perception studies, but also in production studies. For example, vowel space expansion is observed for words with a dense neighborhood [8,9]. VOT enhancement is also observed if a lexical competitor exists [3]. For example, comparing *cod-god* vs. *cop-*gop,* [k] in *cod* has longer VOT than [k] in *cop* because *cod* has a minimal-pair competitor (*god*) while *cop* does not..

However, lexical effects in English apparently occur in and around stressed syllables. Words that have many neighbors tend to be high-frequency words, which are, by and large, monosyllabic and thus stressed. Vowel space is directly affected by stress, and VOT is strongly related to aspiration, both of which are good indicators of stressed syllables.

A question that naturally arises is: what about the lexical effects on phonetic realization in non-stress languages? The phenomenon in question is far under-investigated in such languages. Japanese, which is a non-stress-accent language, has stops with a two-way voicing contrast, voiced or voiceless. The contrast is based on VOT, though without strong voiceless aspiration in any position of the word. It would thus be intriguing to compare lexical effects in VOT between Japanese and English. Another, even more interesting twist about the phonetic realization of voicing contrast in Japanese is that, as was shown in recent studies, VOT in Japanese has been undergoing an ongoing change toward greater overlap between /k/ and /g/, depending on age, region, and gender [10]. As a result, an authentic VOT contrast is being blurred. Lexical effects in voicing contrasts may take other forms than just VOT enhancement in Japanese.

The current study pursues the following three research questions. Q1: Is the phonetic realization of stops affected by lexical competition? That is, is VOT enhanced if there is a minimal-pair competitor? Q2: Do lexical properties other than lexical competition, such as word familiarity, matter? Q3: Given that the voicing contrast is undergoing change, how do lexical effects surface in spoken Japanese? In order to answer these questions, a Japanese speech corpus study was conducted.

2. SPEECH CORPUS

The speech corpus used in this study was the Corpus of Spontaneous Japanese (CSJ) [7]. Approximately 44 hours of speech collected from 201 speakers, along with accompanying annotations, were used for the analyses. In this study, word-initial velar stops were investigated due to the fact that stops in other places of articulation are somewhat irregular in Japanese. Bilabial voiceless stops /p/ are mostly seen in loanwords, mimetics and the coda position in Sino-Japanese words, but not in Yamato words (the core part of the lexicon) [6]. Coronal stops /t/ and /d/ only

appear before non-high vowels except for some recent loanwords.

Sequences of word-initial "<cl>" tag (indicating "closure") followed by "k" were extracted from the phoneme tier labels in the corpus. Merged segments, such as "<cl>,k", and segments with special tags, such as fillers and mispronunciations, were excluded from our analyses. Each [k]-initial token was matched with entries in the Lexical Properties of Japanese database [2] to obtain audio familiarity ratings. On a scale from 1 (not familiar at all) to 7 (highly familiar), tokens with a rating of 5.8 (median of this dataset) or lower were considered low-familiarity words, while tokens with a rating higher than 5.8 were considered high-familiarity words. Tokens that matched multiple homophonic entries in the database were excluded from analysis. In addition, for each /k/-initial word found, the same lexical database was searched to see whether a corresponding /g/-initial word existed as a real word or not. Words that had a corresponding /g/-initial word in the database were considered as having a lexical competitor, while words that did not have a corresponding /g/-initial word were considered as having no competitor.

Table 1 summarizes the number of tokens that were found in the corpus. The tokens were split into four lexical conditions according to two variables: familiarity (low vs. high) and competitor (without vs. with competitor).

Table 1: Number of tokens in each condition, and sample words.

familiarity (audio)	competitor (/g/-initial minimal pair)		total
	without	with	
low (≤ 5.8)	3,561 e.g. /kurabe/	230 e.g. /keta/	3,791
high (>5.8)	2,917 e.g. /kore/	399 /kaisya/	3,316
total	6,478	629	7,107

3. ANALYSES

Linear mixed effects models (R ver.3.4.0, lme4 package ver.1.1-18-1) were used for the analyses. Fixed effects were familiarity (low, high) and competitor (without, with). Dependent variables included durations of the voiceless velar stop and surrounding intervals, i.e. VOT, preceding closure duration, and following vowel duration (for short vowels only).

Various random effects involving speaker and word were added step by step to the model, and likelihood ratio tests were conducted in each step. A best-fitting model that still converged to a solution was adopted. The resulting model included the following random effects: by-speaker and by-word random intercepts, and by-speaker random slopes for familiarity and for competitor.

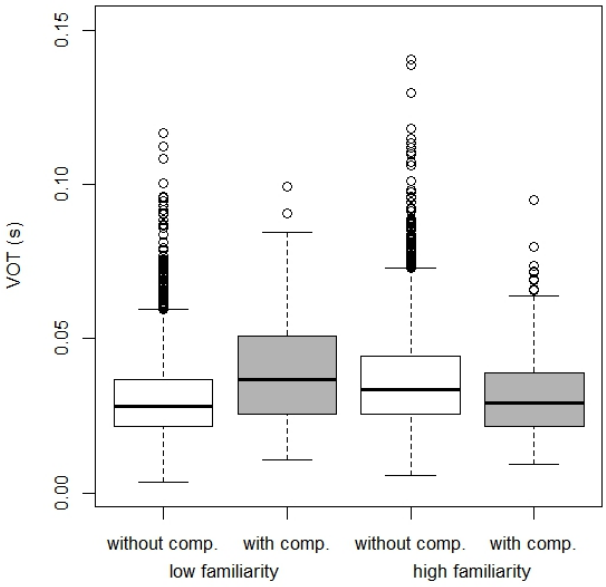
4. RESULTS

4.1. VOT

Figure 1 shows boxplots of the VOT values as a function of the four conditions of the fixed effects. The left half shows results for low-familiarity words, and the right half shows those for high-familiarity words. The white and grey boxes show results for words without a competitor and words with a competitor, respectively.

The analysis showed no significant effect of familiarity, competitor, or their interaction.

Figure 1: Boxplots of the VOT values as a function of the four conditions of the fixed effects.



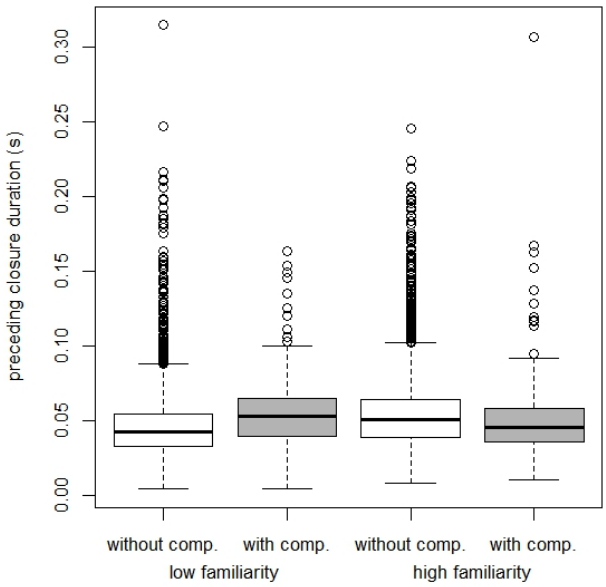
4.2. Preceding closure duration

Figure 2 shows boxplots of the preceding closure duration as a function of the four conditions of the fixed effects. Again, the left half shows results for low familiarity words, and the right half shows those for high familiarity words. The white and grey boxes show results for words without a competitor and words with a competitor, respectively.

The mixed effects model analysis showed a significant effect of familiarity (t = 2.24, p < .05). Preceding closure duration was significantly shorter for low-familiarity words, which had a mean of 0.049s, than for high-familiarity words, which had a mean of 0.055s. The effect of competitor and the

interaction between familiarity and competitor were not significant.

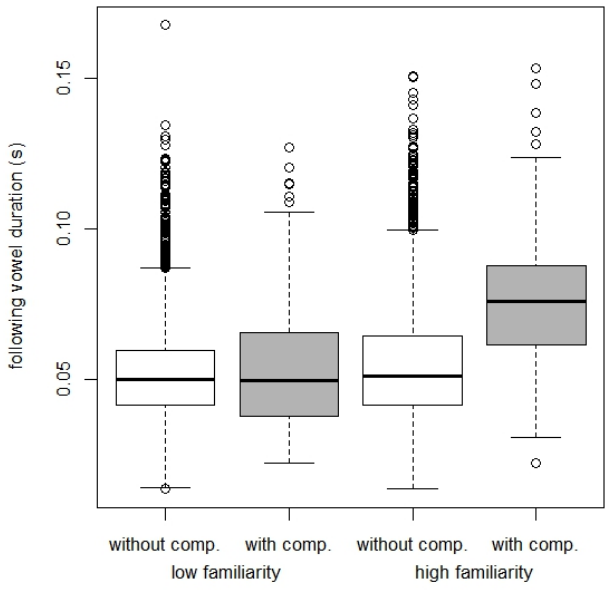
Figure 2: Boxplots of the preceding closure duration values as a function of the four conditions of the fixed effects.



4.3. Following vowel duration

Finally, Figure 3 shows the results for the duration of the following vowel. The analysis here was the same as in Figures 1 and 2. Only phonemically short vowels were included in the analysis.

Figure 3: Boxplots of the following vowel duration values as a function of the four conditions of the fixed effects.



The figure shows a tendency for the following vowel to be longer for words that had a minimal-pair competitor (grey boxes). The mixed effects model showed a significant effect of competitor ($t = 2.02$, $p < .05$). Specifically, vowel duration was significantly longer when a word had a competitor, with a mean duration of 0.069s, than when it did not have a competitor, with a mean of 0.053s. The effect of word familiarity was not significant, nor was the interaction between familiarity and competitor.

5. DISCUSSION

In the present study, three research questions were investigated. Q1 asked whether VOT is enhanced if there is a minimal-pair competitor. The results showed that the means were statistically not significant across conditions, suggesting that the lexical factors did not affect VOT. This was inconsistent with the findings in English [3].

Q2 asked whether lexical properties other than lexical competition, such as word familiarity, affect phonetic output. The results showed a slight though statistically significant effect of word familiarity on the preceding closure duration, such that closure duration was slightly shorter for low-familiarity words than for high-familiarity words. However, the presence of a lexical competitor had no effect on closure duration.

Overall, these results diverge from those found for English [3]; instead, the results are consistent with the interpretation that phonetic consequences of lexical competition may potentially be weaker in non-stress-accent languages, e.g. Japanese. This suggests the importance of examining a wide array of languages in linguistic research.

Q3 asked how lexical effects surface in spoken Japanese, given that the voicing contrast is undergoing change. The results showed that duration of the following vowel was significantly longer when a voicing competitor existed than when it did not. This is a somewhat unexpected finding, but is consistent with the possibility that lexical competition led to a slower, more careful articulation. Also, this finding is parallel to the recent trend in Japanese where voicing contrast in VOT is blurred and transferring to other phonetic dimensions, such as pitch in the following vowel [10].

That the effect of competitor on voicing affected the duration of the following vowel is a substantial and robust effect of about 16 ms even across a huge and uncontrolled dataset. An implication of this finding is that lexical effects are under the influence of moraic organization of syllable structure. That is, a C-V coupling within a mora is tight and co-varying even when the contrast is primarily targeted on the consonant part. Meanwhile, a V-C interaction for

voicing contrast is clearly attested in English. When the coda consonant is voiced, the preceding vowel is about 50% longer than a vowel before a voiceless consonant as in the minimal pair *bit-bid* [5] (also known as pre-fortis clipping). Not surprisingly, this effect is not stable in adults' speech in Japanese even though Japanese infants show the effect until the age of five [1]. In other words, the V-C interaction across a mora boundary observed in infancy is overridden by the establishment of moraic organization in Japanese prosody.

6. CONCLUSION

In the present research, the duration of intervals around the burst of word initial velar stops in the Corpus of Spontaneous Japanese was investigated to see if lexical factors, such as competitors in a minimal pair and word familiarity, affect the phonetic realization of voicing contrast in Japanese. There was no significant effect of such factors on VOT, a supposedly primary cue of voicing, however. Nevertheless, there were significant effects of competitors on the duration of the following vowel and word familiarity on the closure duration. This suggests that lexical effects do exist in Japanese, but in a radically different manner from English in which the contrast enhancement effect was attested in VOT.

The lexical effect of voicing in the following vowel found in the present study can be primarily interpreted as a slower, more careful articulation in the presence of a competitor to enhance the contrast. However, language-particular prosodic structure may restrict the realization of contrast enhancement in a specific direction: i.e., a C-V coupling within a mora may be tight to enforce the vowel prolongation even when the enhancement is supposedly taking place on the consonant part. Though this hypothesis needs a more careful investigation in the future study, we can draw a very general implication: Japanese, a language that is phonemically, prosodically and typologically different from English, is a useful test-bed for analyzing cross-language generalizability of the phonetic realization of phonological organization and contrast.

There are several limitations to this study that need to be addressed. One problem was that there was a very large by-item variability in the VOT values. This was shown by the fact that the effect of competitor, which was significant in initial analyses, was eliminated when the random effect of word was included in the model. This is not surprising since this was a corpus study, and as such, it was difficult to control for covariates such as phonetic context, within-utterance position, speaking rate, to name a few. It can be pointed out that including everything in a statistical model is a possible solution. However, there are not enough "with-competitor" tokens in the corpus. Speaker variables, such as age, gender, and region can also be included in our statistical analysis. Log-transformation of the dependent variable is another option to circumvent the problem of skewed distribution. All these statistical techniques and applications will be our next task.

Another problem was that words have many neighbors, not just voicing competitors. The words analyzed in the present study differ from other words not only in the voicing of the initial consonant, but also in place and manner, and also with respect to other segments in the word. Competition with these neighbors may also affect phonetic realization of these words. Therefore, a well-controlled experiment with a direct comparison of minimal pairs, such as /kama/-/gama/, is needed.

7. ACKNOWLEDGEMENTS

This work was supported by JSPS KAKENHI Grant Numbers 16K02646, 18K00549, and 18K00662.

8. REFERENCES

[1] Kitahara, M. and Yoneyama, K. 2014. Voicing effect on vowel duration: Corpus analyses of Japanese infants and adults, and production data of English learners, *Journal of the Phonetic Society of Japan* 18, 30-39.

[2] Amano, S., Kondo, T. 1999. *Lexical Properties of Japanese*, Tokyo: Sanseido.

[3] Goldrick, M., Vaughn, C., Murphy, A. 2013. The effects of lexical neighbors on stop consonant articulation. *JASA* 134 http://dx.doi.org/10.1121/1.4812821

[4] Greenberg, J. H., Jenkins, J. J. 1964. Studies in the psychological correlates of the sound system of American English. *Word* 20, 157–177.

[5] House, A. S. 1961 On vowel duration in English, *J. Acoust Soc. America*, 33, 1174-1178.

[6] Ito, J., Mester, A. 1995. The core-periphery structure of the lexicon and constraints on reranking. In Beckman, J., Urbanczyk, S., Walsh, L., eds. *University of Massachusetts Occasional Papers in Linguistics [UMOP]* Vol. 18: Papers in Optimality Theory. University of Massachusetts, Amherst: GLSA. 181-209.

[7] Maekawa, K. 2003. Corpus of Spontaneous Japanese: Its design and evaluation. Proc. of SSPR2003, Tokyo.

[8] Munson, B., Solomon, N.P. 2004. The effect of phonological neighborhood density on vowel articulation. *Journal of Speech Language and Hearing Research* 47, 1048-1058.

[9] Scarborough, R.A. 2010. Lexical and contextual predictability: Confluent effects on the production of vowels, *Laboratory Phonology* 10, 557-586.

[10] Takada, M., Kong, E.J., Yoneyama, K., Beckman, M.E. 2013. Loss of prevoicing in modern Japanese /g, d, b/. *Proc. 18th ICPhS* Glasgow, paper number 0873.