Japanese Velar Allophones Revisited: a Quantitative Analysis Based on the Speech Production Experiments

SHIN-ICHIRO SANO
Dokkyo University

1. Introduction
This research reexamines the voiced velar allophones /g/ in Japanese through the quantitative analysis based on the speech production experiments, taking phonological factors governing the distribution of these allophones into account. It has been traditionally assumed that velars in Japanese realize as either plosives [ɡ] or nasals [ŋ] (Kindaichi 1942; Otsu 1980; Kato 1983; Shibatani 1990; Hibiya 1995, 1999; Inoue 1998; Okada 1999; Kindaichi and Akinaga 2001; Inoue and Yarimizu 2002; Vance 1987, 2008, among others). The distribution of velar allophones has been assumed to be governed by the contextual factors. The fundamental restriction concerns with the position within a word: velars realize as plosives in word-initial

*I would like to thank Ross Krekoski, Ichiro Yuhara, Timothy Vance, Natsuko Tsujimura, audiences at the 19th conference, and reviewers for their valuable comments. Special thanks also go to the members of Sophia Phonetics Lab. and Frank Scott Howell for their help and support on this and an earlier version of this paper. Any remaining faults are, of course, mine.

1 In the following, I refer to voiced velars (allophones) as simply velars (allophones).
positions as in *gakuhi* ‘tuition’ (Figure 1); on the other hand, in word-
internal positions as in *dokuga* ‘venom fang,’ velars realize as nasals (Fig-
ure 2).\(^2\)

![Figure 1. Waveform and spectrogram of \[\gamma\] in gakuhi (speaker #03)](image1)

![Figure 2. Waveform and spectrogram of \[\eta\] in dokuga (speaker #01)](image2)

The positional restriction interacts with some other factors such as the lexi-
cal strata (e.g. Sino-Japanese, Yamato Japanese, or loanwords), and the
word-internal structure in compounds (Kindaichi 1942; Vance 1987, among
others).\(^3\) The intricate interactions of contextual factors define the variable

\(^2\) For the details of the recording and the phonetic analysis, see Section 2.

\(^3\) At this point, I consider the reason for the idiosyncratic distribution of velar allophones in
onomatopoeic words. Japanese onomatopoeic words involve reduplication as in *garagara, mogmog*. The allophonic realization of velars in these words obeys distinct rules: 1) in initial
positions of each constituent (a base and a reduplicant), velars categorically realize as plosives
distribution of velar allophones. However, a few works point out that velars can realize as fricatives [ɣ] as in Figure 3 (spirantization, Kintaichi 1942; Kamei 1956; Hattori 1957).

![Figure 3. Waveform and spectrogram of [ɣ] in eigo ‘English’ (speaker #02)](image)

Furthermore, the distribution of velar allophones shows some variation and change. In terms of the dialectal difference, in Tohoku dialect nasals are remarkable; on the other hand, in Kansai dialect, velars categorically realize as plosives, and in Kanto dialect the situation is intermediary (Vance 1987, among others). Thus, the dialectal difference of velar allophones is characterized as in Tohoku dialect > Kanto dialect > Kansai dialect, with respect to nasals. As for the change, the decline of velar nasalization has been reported, namely, cases have been increasing where velars in word-internal positions realize as plosives instead of supposed nasals, in violation of the aforementioned restrictions (Kindaichi 1942; Hibiya 1995).

Specifically, Kindaichi (1942) claims that the rate of velar nasals has been declining as in (ガガガガ), even though the initial position of the second constituent is a word-internal one; 2) velars categorically realize as nasals, only if the velars appear in final positions of each constituent as in [もぎもぎ] (Vance 2008). Based on the assumption that the Japanese reduplication involves the process where the first constituent functions as the base and the second one functions as the reduplicant obtained by copying the base, I argue that in the former case the reduplicant is obtained by copying the base which includes velars in word-initial positions ([ガガガガ]); it follows that the reduplicant also includes word-initial velars and all velars realize as plosives, even though the velars in the reduplicant themselves position word-internally; on the other hand, in the latter case the reduplicant is obtained by copying the base which includes velars in word-internal positions ([もぎもぎ]); it follows that the reduplicant also includes word-internal velars and all velars realize as nasals. Thus, the difference between the former case and the latter one can be attributed to whether the process involves the copying of word-initial velars or of word-internal velars.

---

As in ([ガガガガ]), even though the initial position of the second constituent is a word-internal one; 2) velars categorically realize as nasals, only if the velars appear in final positions of each constituent as in [もぎもぎ] (Vance 2008). Based on the assumption that the Japanese reduplication involves the process where the first constituent functions as the base and the second one functions as the reduplicant obtained by copying the base, I argue that in the former case the reduplicant is obtained by copying the base which includes velars in word-initial positions ([ガガガガ]); it follows that the reduplicant also includes word-initial velars and all velars realize as plosives, even though the velars in the reduplicant themselves position word-internally; on the other hand, in the latter case the reduplicant is obtained by copying the base which includes velars in word-internal positions ([もぎもぎ]); it follows that the reduplicant also includes word-internal velars and all velars realize as nasals. Thus, the difference between the former case and the latter one can be attributed to whether the process involves the copying of word-initial velars or of word-internal velars.
starting with speakers born in the 1910s;\textsuperscript{4} Hibiya (1995) claims that 48\% of all the velars in word-internal positions realized as nasals in the data collected in 1986.

However, as we saw above, the in-depth analysis of the data has been insufficient, phonetic/phonological aspects of the distribution of velar allophones have been underresearched, and crucially the velar fricatives has been overlooked, as the previous studies had mainly been based on the “word” as their units of analysis. The goals of the present research, therefore, include: 1) to challenge the claims of previous studies such as the traditional plosive/nasal dichotomy and the decline of velarnasalization; 2) to examine the effects of factors, which govern the distribution of velar allophones. Specifically, the analysis focuses on the following points: the distributions 1) in word-initial positions, 2) in post-nasal positions, 3) in sequential contexts; and 4) the effects of following vowels; 5) the dialectal difference. Throughout the analysis, I assume the trichotomy among plosives, nasals and fricatives following the claim of spirantization.\textsuperscript{5}

This paper is organized as follows: Section 2 introduces the details of the experiments; in Section 3, I summarize the results; in Section 4 I present the analysis and the discussion; Section 5 concludes the discussion.

2. Method

2.1. Target words

In the speech production experiments, I selected target words exclusively from Sino-Japanese words with three morae length for the purpose of eliminating the effects of word length and of lexical strata. In addition, in Sino-Japanese words the distribution of velars is relatively free, allowing us to set up various contexts. Specifically, I arranged target words focusing on the following factors: phonological contexts such as word-initial velars, post-nasal velars, and sequential velars; the types of following vowels. The target words include the following six phonological contexts:

- (1) $\overline{u}g$ (no carrier sentence) e.g. gokushi ‘tuition’
- (2) $\overline{u}g$ e.g. gimi ‘review’
- (3) Vg e.g. dokuga ‘venom fang’
- (4) Ng e.g. giga ‘galaxy’
- (5) VgVg e.g. vigo-ga ‘English-case particle’
- (6) NgVg e.g. rongo-ga ‘Analects-case particle’

\textsuperscript{4} Kindaichi (1942) does not conduct the recordings of audio data.

In the present data, some velar allophones, which show the characteristics similar to voiced velar approximants, were observed. I include these allophones into fricatives, instead of establishing an extra independent category.
Velars, in themselves, appear in word-initial positions both in (1) and (2). In (1), however, target words are not embedded in carrier sentences (kare-wa ‘……’ to itta “He said ‘ …… ’,” yielding velars in utterance-initial positions; on the other hand, in (2) target words are embedded in carrier sentences, yielding velars in word (prosodic word, accentual phrase) -initial positions. Every target word is presented as embedded in carrier sentences except for (1). In (3) velars appear in post-vocalic positions. In (4) velars appear in post-nasal positions. (5) and (6) include the sequential contexts where syllables involving velars are adjacent. (5) involves two post-vocalic velars; on the other hand, (6) involves post-nasal velars followed by post-vocalic velars. Every sequential context includes the case-particle ga.

For each context listed above, I arranged extra five contexts with respect to the types of following vowels (/gi/, /ge/, /ga/, /go/, /gu/).

2.2. Participants
I randomly sampled three participants according to the following criteria: 1) age: under 25; 2) gender: male/female; 3) hometown: Tohoku/Kanto/Kansai areas; 4) experience abroad: less than one month. Specifically, in order to examine the progress of velar nasalization, I focused on younger speakers; for the examination of dialectal differences, I sampled one speaker from each of these areas; and to eliminate the influence of foreign languages, I limit the experience abroad within the short time-period. In the prior questionnaire, no participants reported auditory impairments.

2.3. Recording
The recordings have been conducted in the Sophia phonetics lab. attached soundproof chamber for each participant individually. The recording conditions are as follows: recorder: SONY linear PCM recorder; Microphone: SONY ECM-959DT (directionality: 90°); Sampling frequency: 48 kHz; Quantization: 16bit; monaural digitization. In the recordings, I randomly presented target words and distracters embedded in carrier sentences one by one; participants read the whole sentence according to the presented script. I recorded at least five tokens for each target word/distracter.

2.4. Segmentation
The recorded audio data were in turn analyzed by Praat (version. 4.5.08, Boersma and Weenink 2006) Spectrogram settings were as follows: View range: 0 ~ 5000Hz; Window length: 0.005s; Dynamic range: 50dB. I ex-

6 The attributes of participants are shown in the Appendix. In the pilot study, I conducted the recording as well as the analysis of audio data of one extra participant for each gender.

7 Although in some cases I recorded more than six tokens for a certain word, only five out of six tokens were subjected to the analysis.
tracted the parts corresponding to velar allophones from the audio data and classified these into plosives, fricatives, and nasals. The criterion for the classification of velar allophones is schematized below.

\[
\begin{align*}
\text{Obvious burst} & \quad \text{Yes} \quad \rightarrow \quad \text{[g]} \\
& \quad \text{No} \quad \rightarrow \quad \text{Nasal formant} \\
& \quad \text{Yes} \quad \rightarrow \quad \text{[ŋ]} \\
& \quad \text{No} \quad \rightarrow \quad \text{[y]}
\end{align*}
\]

Figure 4. Criterion for the classification of velar allophones

Firstly, if an obvious burst (closure and release) is observed with respect to the waveform (e.g. Figure 1), the velar allophone is categorized as a plosive; if no obvious burst is observed, then I refer to the nasal formant; if a nasal formant is observed (e.g. Figure 2), the velar allophone is categorized as a nasal; if no nasal formant is observed (e.g. Figure 3), then the velar is categorized as a fricative.\(^8\)

3. Results

As mentioned above, I arranged six phonological contexts and five contexts for the types of following vowels. In addition, each target word in sequential contexts includes two velars. Thus, the total tokens I focus on amount to 600 (6 phonological context \(\times\) 5 vowels \(\times\) 5 tokens \(\times\) 3 speakers + 2 phonological context \(\times\) 5 vowels \(\times\) 5 tokens \(\times\) 3 speakers). The overall distribution of velar allophones is summarized in Table 1.

<table>
<thead>
<tr>
<th>Table 1. Distribution of velar allophones</th>
</tr>
</thead>
<tbody>
<tr>
<td>frequency</td>
</tr>
<tr>
<td>277</td>
</tr>
<tr>
<td>269</td>
</tr>
<tr>
<td>54</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

As Table 1 shows, the frequency (269) as well as the probability (44.83\%) of fricatives are almost the same as those of plosives (frequency: 277; probability: 46.17\%); on the other hand, the frequency (54) and the probability (9\%) of nasals are extremely low. The result shows that velar fricatives, which have been overlooked in previous works, are frequently observed in

\(^8\) The discrimination between velar nasals and velar fricatives in pos-nasal positions (Ng) involved some difficulties.
actual utterance and that the decline of velarnasalization has been accelerated, considering the younger age of participants.

Next, I show the distribution of velar allophones by six phonological contexts arranged in the target words.

![Figure 5. Distribution of velar allophones by phonological contexts](image)

In Figure 5, “NgVg\_1st” and “VgVg\_1st” represents the preceding velars in sequential contexts (e.g. rongo-ga, eigo-ga), and “NgVg\_2nd” and “VgVg\_2nd” represents the following velars (e.g. rongo-ga, eigo-ga).

As Figure 5 shows, the probability of plosives is 100% in utterance-initial positions and 80% in word-initial positions. In word-initial positions, plosives are shown to be predominant; on the contrary, no nasals are observed. In postnasal positions nasals show a relatively higher probability (Ng: 35%; NgVg\_1st: 21%) compared with other positions. In sequential contexts, we can observe the higher probability of fricatives (NgVg\_2nd: 75%; VgVg\_1st: 64%; VgVg\_2nd: 85%).

Based on the results, I make a rough generalization of the distribution of velar allophones with respect to the phonological context: 1) word-initial positions – [q]; 2) post-nasal positions – [ŋ]; 3) sequential contexts – [y].

4. Analysis

In this section, I will conduct the detailed analyses as well as the discussions for each context.

---

9 Every distribution in the following figures was shown to be statistically significant, although I omit the details due to space limitations.
4.1. Word-initial position
Firstly, I examine the effect of word-initial positions. Although previous studies claim that velars realize as plosives in word-initial positions, it is unlikely that words are uttered separately in spontaneous speeches; instead, words are usually surrounded by preceding and following constituents in a continuous stream of utterances. Given that Japanese basically takes the CV syllable structure, it follows that velars are surrounded by preceding and following vowels (CVgV) even in word-initial positions, and velars can realize as fricatives according to the aforementioned spirantization. To verify this hypothesis, I examine the distribution of velar allophones by word-initial and other positions. The result is shown below.

As shown in Figure 6, in word-initial positions the probabilities of plosives and of fricatives are 90% and 10%, respectively, and no nasals are observed, demonstrating that plosives are predominant in word-initial positions. In other positions, on the other hand, the probability of plosives is 32%, that of fricatives is 56%, and that of nasals is 12%, showing that fricatives are more frequent and plosives are less frequent than in word-initial positions. Thus, the word-initial positions were shown to have a strong impact on the distribution of velar allophones; however, the claim of spirantization was also verified in actual utterances, considering the fact that velar fricatives, albeit only slightly, were attested.

Next, I examine the distribution within word-initial positions. Specifically, I examine the difference between utterance-initial positions ([U g]) and word-initial positions ([ω g]) with respect to their effects on the distribution of velar allophones. I illustrate the distribution of velar allophones by utterance-initial and word-initial positions in Figure 7.

As Figure 7 shows, every velar categorically realizes as plosives in utterance-initial positions; on the other hand, in word-initial positions the probability of plosives is 80% and that of fricatives is 20%, and no nasals are observed. The result slightly differs from the claims of previous studies:
although plosives are predominant, not a few fricatives are observed in word-initial positions.

![Figure 7. Distribution of velar allophones in word-initial position](image)

Thus, we can argue that the difference in utterance-initial and word-initial positions affects the distribution of velar allophones, and that velars can realize as other allophones than plosives due to some phonological factors even if they are in the same positions within a “word.” The result further shows that in word-initial positions velar allophones are not categorically restricted to plosives; rather, the distribution should be characterized in a gradient manner; for example, plosives are more frequent than fricatives in word-initial positions.

At this point, I consider the reason for the different distributions in utterance-initial and word-initial positions in terms of the difference in levels of strengthening in prosodic structure: initial segments of each prosodic category undergo the strengthening; the degree of strengthening depends on the level of the category in prosodic structure: the higher the level, the stronger the strengthening (e.g. utterance > intermediate phrase > accentual phrase > prosodic word) (Keating 2006, among others). In the present case, the utterance-initial positions are higher, while word-initial positions, which are categorized as accentual phrases or prosodic words, are lower with respect to the levels in prosodic structure. This gives more chance for word-initial positions to be spirantized, where weaker strengthening is expected.

4.2. Post-nasal position

I turn now to the examination of the distributions in post-nasal positions. In post-nasal positions, it can be predicted that velars are likely to realize as nasals, being affected by the preceding nasals. To verify this hypothesis, I examine the distribution of velar allophones by post-nasal and other positions. The result is shown in Figure 8.

Focusing on the distribution of nasals, we can recognize that nasals show a higher probability in post-nasal positions (28%) than in other positions (3%). In post-nasal positions, velars preferentially realize as nasals compared with other positions, in support of the hypothesis.
This process can be thought of as a kind of progressive assimilation. All in all, however, the probability of nasals is lowest in either context compared with that of plosives and of fricatives, showing that the higher likelihood of nasalization in post-nasal positions is suppressed by the effect of the decline of velarnasalization. The interaction among factors is demonstrated to play a crucial role in defining the distribution.

4.3. Sequential context

In this part, I examine the distribution in sequential contexts. In sequential contexts, it can be hypothesized that time for tongue movement is insufficient, resulting in the halfway closure, and more fricatives are observed. To verify this hypothesis, I examine the distribution of velar allophones by sequential and single contexts. The result is shown below.

As shown in Figure 9, in single contexts the probability of fricatives is 23% and plosives show the highest probability (67%); on the other hand, in sequential contexts fricatives show the highest probability (66%). The result shows that in sequential contexts velars are more likely to realize as fricatives, as hypothesized.

Next, I analyze the distribution within sequential contexts. Firstly, which one is more likely to realize as fricatives, preceding velars or following velars? Preceding velars cannot be affected by other preceding velars as there are no other preceding velars; on the other hand, following velars may
well be affected by preceding velars, yielding more fricatives. I illustrate the distribution of velar allophones by preceding/following distinction.

As Figure 10 shows, the probability of fricatives is extremely high in following velars (80%), as predicted; however, the high probability of fricatives is also observed in preceding velars (53%). The result shows that in sequential contexts following velars are more likely to realize as fricatives, and that the preceding and the following velars are interacting with each other. Note that in the present analysis the following velars are all included in case particles, and most of the following velars realize as fricatives, contrary to the claims of previous studies that velars in case particles categorically realize as nasals (Kindaichi and Akinaga 2001, and others). The result suggests the importance of phonological factors rather than grammatical factors.

Secondly, I take a closer look at the relationship between preceding and following velars. Specifically, I examine whether preceding and following velars in a single token match or mismatch with respect to the allophonic realization. Given the result that the distributions of preceding velars and of following ones are different, I hypothesize that preceding and following velars mismatch. I verify the hypothesis by cross-tabulation.

Table 2. Cross-tabulation of the distributions of preceding and of following

<table>
<thead>
<tr>
<th></th>
<th>g</th>
<th>γ</th>
<th>η</th>
</tr>
</thead>
<tbody>
<tr>
<td>g</td>
<td>15 (60%)</td>
<td>6 (24%)</td>
<td>4 (16%)</td>
</tr>
<tr>
<td>γ</td>
<td>37 (31%)</td>
<td>73 (61%)</td>
<td>10 (8%)</td>
</tr>
<tr>
<td>η</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>5 (100%)</td>
</tr>
<tr>
<td>Total</td>
<td>52</td>
<td>79</td>
<td>19</td>
</tr>
</tbody>
</table>

χ² = 47.03, d.f. = 4, p = 1.504e-09, fisher’s exact test: p = 6.664e-07
In Table 2, cases where preceding and following velars match show a higher probability (shaded cells). Furthermore, putting all tokens together, the probability of “match” constitutes 62% of all cases (“mismatch”: 38%), showing that preceding and following velars in sequential contexts are more likely to match, contrary to the hypothesis. Note that in sequential contexts velars are not adjacent as there is an intervening vowel as in *eigoga*. Therefore, I argue that the matching in the present case is an instance of assimilations in nonlocal contexts, in support of the Long Distance Consonant Agreement (LDCA: Rose and Walker 2003), which assumes the agreement of consonantal features in nonlocal contexts.

4.4. Effects of following vowels

A number of phonetic/phonological researches to date point out that consonants undergo some changes with respect to their features, because they are affected by following vowels (Keating et al. 1994; Ladefoged 2006, among others). If this is on the right track, the distribution of velar allophones should vary according to the types of following vowels. I examine the distribution of velar allophones for five vowels.

![Figure 11. Distribution of velar allophones by types of following vowels](image)

As shown in Figure 11, the probability of plosives is higher before /i/ and /e/ vowels; on the other hand, the probabilities of fricatives and nasals are higher before /a/ and /o/ vowels. The result shows that the distribution of velar allophones is affected by types of following velars as expected: specifically, plosives are more compatible with high and front vowels, while fricatives and nasals are more compatible with low and back vowels. This can be attributed to articulatory reasons: to articulate velars followed by low and back vowels, a relatively larger tongue movement is required, compared with high and front counterparts, resulting in the halfway-closure, and velars are likely to realize as fricatives or nasals, rather than plosives.
4.5. Dialectal difference

Finally, I examine the distribution in terms of the dialectal difference. As mentioned above, the dialectal difference of velar allophones is characterized as in Tohoku dialect > Kanto dialect > Kansai dialect, with respect to nasals. Even with the decline of velar nasalization, is the characterization of dialectal difference still valid? If so, we would obtain the similar distribution as above even in the present analysis, which focuses on utterances of younger speakers. For the examination, I classified velar allophones according to the hometown of speakers. The result is shown below.

As Figure 12 shows, with respect to nasals Tohoku shows remarkably higher probability (20%), and nasals are frequent in the following order: Tohoku > Kanto > Kansai, showing that the characterization of dialectal difference remains valid. Although no significant difference is observed with respect to plosives, the following tendency is observed concerning fricatives: fricatives are most frequent in Kansai, least in Tohoku, and Kanto is in between. This order is the exact opposite of the one in nasals.

5. Conclusion

In this paper, I presented the analysis of Japanese velar allophones. Specifically, I challenged the traditional categorization of velar allophones, and I examined the distribution of these allophones taking phonetic/phonological factors into account. The examinations brought forth the following findings: 1) the decline of velar nasalization has been accelerated; 2) word-initial velars can realize as fricatives, although plosives are predominant (contra Kindaichi 1942; the distribution is governed by the prosodic structure (cf. Keating 2006)); 3) post-nasal velars preferentially realize as nasals; 4) velars in sequential contexts preferentially realize as fricatives; 5) progressive assimilation is observed in post-nasal positions and sequential contexts (LDCA, Rose and Walker 2004); 6) velars preceding front and high vowels preferentially realize as plosives, while those preceding low and
back vowels are more likely to realize as fricatives or nasals; 7) the characterization of dialectal difference is still valid (nasals: Tohoku dialect > Tokyo dialect > Kansai dialect). Incidentally, the distribution of velar alloglyphs in Japanese shows some similarity with those in Spanish with the plosive/fricative dichotomy, except for nasals (Japanese: utterance-initial positions: plosives, post-nasal positions: nasals, other contexts: fricatives; Spanish: utterance-initial/post-nasal positions: plosives, other contexts: fricatives) (cf. Hualde 2005).

I propose that the traditional plosive/nasal dichotomy is insufficient to figure out the nature of the velar alloglyphs in Japanese; instead, the velar alloglyphs need to be discussed in terms of the trichotomy with fricatives; the distribution of velar alloglyphs is defined by the interaction between language changes and phonological factors. Furthermore, it can even be predicted that nasal will disappear; the velar alloglyphs will be comprised only of plosives and fricatives.

Appendix: attributes of participants

<table>
<thead>
<tr>
<th>Speaker #</th>
<th>age</th>
<th>gender</th>
<th>hometown</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>19</td>
<td>male</td>
<td>Iwate</td>
</tr>
<tr>
<td>02</td>
<td>25</td>
<td>female</td>
<td>Kanagawa</td>
</tr>
<tr>
<td>03</td>
<td>22</td>
<td>male</td>
<td>Nara</td>
</tr>
</tbody>
</table>

References


