

# Japanese Vowel Deletion Occurs in Words in Citation Form

Alexander Kilpatrick<sup>1</sup>, Rikke L. Bundgaard-Nielsen<sup>2</sup>, Brett J. Baker<sup>1</sup>

<sup>1</sup>University of Melbourne, Australia

<sup>2</sup>MARCS Institute for Brain, Behaviour and Development,  
Western Sydney University, Australia

akilpatrick@student.unimelb.edu.au, rikkellou@gmail.com, bjbaker@unimelb.edu.au

## Abstract

Japanese vowels have allophonic reduced variants, including shortened, devoiced and deleted instances. This kind of linguistic behaviour is commonly associated with rapid or casual speech. However, the present report demonstrates that vowel deletion also occurs when Japanese speakers produce Japanese words in citation form, and not just in rapid or casual speech. We propose that deletion is more likely to occur in high-frequency lexical items, specifically in three regularly occurring suffixes.

**Index Terms:** Japanese, vowel deletion, elision

## 1. Introduction

Cross-linguistically, segmental reduction is typically associated with fast speech. In Japanese, however, vowel reduction—shortening, devoicing, and deletion—is said to occur regularly for the high vowels, /i/ and /u/, even at a normal speaking rate [1]. Shortening, devoicing, and deletion behaviours are generally regarded as features of a single phenomenon, referred to as ‘vowel devoicing’ or ‘reduction’. Reduced vowels are prominent in many Japanese dialects though they characterise the dialects of Kanto and Kyushu in particular [2]. The existing literature suggests that devoicing is more likely to occur when a vowel is preceded by a stop while deletion is more likely to occur when a vowel is preceded by a fricative [3]. [1] argues that it is impossible to determine whether certain vowels are devoiced or deleted, particularly when they occur after fricatives, even by examining the acoustic signal.

In the following, we present a study of the production of Japanese words in citation form and argue that deletion is a feature of citation form and is more likely to be elicited by specific words that occur in high frequency in Japanese discourse.

## 2. Background

Cross-linguistically, vowel production is typically characterised by regular vocal fold vibration (voicing). However, Japanese exhibits a range of devoicing patterns that challenge this observation. Japanese high vowels are regularly devoiced when they are both preceded and followed by unvoiced consonants or when they appear the end of a word and are preceded by an unvoiced consonant. Japanese vowels are also said to undergo deletion in certain environments, although there is some contention regarding this. Those reports [3] that argue that deletion is occurring, suggest that vowels are deleted when they follow fricatives and devoiced when they follow stops. Other reports [2] treat deletion as an extreme variant of devoicing, suggesting that reduction is a continuum, at one end

of which, vowels are voiced and maintain regular duration and deleted at the other, with devoicing occurring at a midpoint.

It is unlikely that Japanese vowel reduction is a recent development. From as early as the 1600s, non-Japanese scholars have been remarking that Japanese high vowels are sometimes inaudible [2]. However, devoicing patterns in Japanese vary across regional dialects: Systematic reduction is a feature of the Kanto dialect but far less frequent in the Kansai dialect [4]. Indeed, cities in East Japan have higher devoicing rates, 68% in Nagoya, 56% in Tokyo, and 56% in Sendai, than cities in the Kansai and Shikoku regions of central Japan, 32% in Osaka, 28% in Okayama, and 18% in Kochi, however, further west, these rates increase to those levels found in Eastern cities, 57% in Naha and 53% in Kumamoto [5].

Vowel reduction behaviour, including devoicing, primarily occurs with the high vowels, /i/ and /u/ [6]. Reduction may, however, also occur with other vowels but this is irregular, arising only in specific lexical items. A frequently cited example is the devoicing of the first /o/ in the Japanese word *kokoro* [heart], as is the first /a/ in *katana* [sword] [3]. Irrespective of the vowel, reduction occurs in two environments; when a high vowel is both preceded and followed by a voiceless consonant, and when a high vowel occurs word finally and is preceded by a voiceless consonant.

Japanese verbs largely follow predictable conjugation rules, though there are a small number of exceptions. The polite, present tense suffix—*ます/masu/*—conjugates by deleting the word-final mora (or final two mora in exceptions) whereas imperative and past tense suffixes—*て/te/* and *た/ta/* respectively—conjugate differently, depending on the word-final mora. For instance, verbs ending in *す/su/* or *する/suru/* conjugate by deleting the word-final mora and adding *して/ite/* or *した/ita/* but conjugate differently for verbs ending in alternative mora.

## 3. Method

### 3.1. Participants

Twelve native Japanese speakers (10 female; 2 male) living in Melbourne, Australia, were recruited for this study. Seven of the participants were expatriates, having lived in Australia for a minimum of five years. Expatriate participant ages ranged from 27-42. The remaining five participants were international students who had lived in Australia for less than a year and were studying English as a Foreign Language at La Trobe University. International Student participants ranged in age from 18-20.

Participants were recruited by word of mouth. Four of the participants indicated that they were native to districts of Japan where the regional dialects are characterised by lack of vowel reduction (see above). These included the Kansai and Shikoku regions. The remaining eight participants indicated that they

were native to areas where vowel reduction is a common feature of the regional dialects, these included Tokyo, Yamagata, Nagoya and Kyushu.

### 3.2. Recordings

Recordings were conducted in quiet rooms in Melbourne, using a Zoom H4n recording device with a sampling depth of 24kb/sec and a sample rate of 44.1kHz. Participants were asked to read from a pseudorandomised list of thirty words in Hiragana script. Participants received no coaching in how they were to pronounce the list of words, and we provided no feedback on their productions.

The word list contained lexical selected to induce deletion or vowel reduction, on the basis of the existing literature [2], [3]. Indeed, we hypothesised that participants would delete vowels in the following contexts:

- /i/ in the imperative verb suffix in e.g. して /ʃite/ [IMP],
- /i/ in the past tense verb suffix in e.g. した /ʃita/ [PST],
- /u/ in the polite, present tense verb suffix in e.g. ます /masu/ [PRS].

We included a number of words containing each of these suffixes (see Table 1 for the list). Amongst the words, we included the suffix by itself to determine whether there was a difference between the suffix presented in isolation and when attached to a verb. We also included three tokens often reported as examples of low vowels that frequently undergo devoicing, ころ /koro/ [heart] [1], ほこり /hokori/ [pride] [7], and はか /haka/ [grave] [8]. Two words, した /ʃita/ [fern] and して /ʃide/ [paper streamer], were included to test the influence of having a voiced consonant follow the vowel in question. Four words, からす /karasu/ [crow], まし /masi/ [better], すき /suki/ [like], and だいすき /daisuki/ [really like] were included to test whether deletion would occur in similar but not identical environments to the suffixes listed above. Seven foils were chosen at random from a Japanese dictionary. We included a single instance of word with the post-stop environment: もく /moku/ [wood], to test the difference (if any) between fricatives and stops. Speakers produced three repetitions of each of the 30 items, resulting in a total of 90 tokens for analysis per participant, and a total of 1080 tokens included in the present study.

### 3.3. Predictions

We predict that speakers will delete high vowels when they are preceded by voiceless fricatives and either occur word finally, or followed by voiceless consonants. Furthermore, we predict that deletion will occur more frequently in these environments when they occur in one of the three suffixes, して /ʃite/ [IMP], した /ʃita/ [PST] and ます /masu/ [PRS], and that this is likely due to their high frequency of occurrence.

### 3.4. Analysis

Elicitations were categorised on the basis of visual inspection of spectrograms. Because there is controversy regarding the existence of deleted vowels in Japanese, we provide the following detailed explanation as to how we categorised participant behaviour. Voiced vowels are characterised by voicing bars—which are indicative of vocal fold vibration—and formant resonance in spectrograms. Unvoiced vowels do not yield voicing bars but do generate formant resonance. Figure 1 shows three instances of the elicitation of the Japanese word, だいすき /daisuki/ [really like], produced by one of our participants, in succession. Across all participants, we observed that the /u/ in this particular environment was frequently deleted. In the first iteration, there are clear voice bars; there is considerable energy in the lower frequencies; and a substantial increase in intensity. While this is a shortened, low intensity vowel, the voice bars indicate that the vocal folds are vibrating and hence that the vowel is voiced. We provide the second iteration as an example of a devoiced /u/. Here, we observe a lack of voice bars and a decrease in both low frequency energy and intensity when compared to the first iteration. There is a small but very clear increase in intensity in this devoiced example. In the third example, there is no vowel. There is no transition from the fricative until the sudden reduction of energy which indicates closure for the following stop. The intensity shows a steady weakening with no instances of incline as can be seen in the second example. It should be noted here that the second example in Figure 1 was the only elicitation in which we observed a devoiced vowel following a fricative.

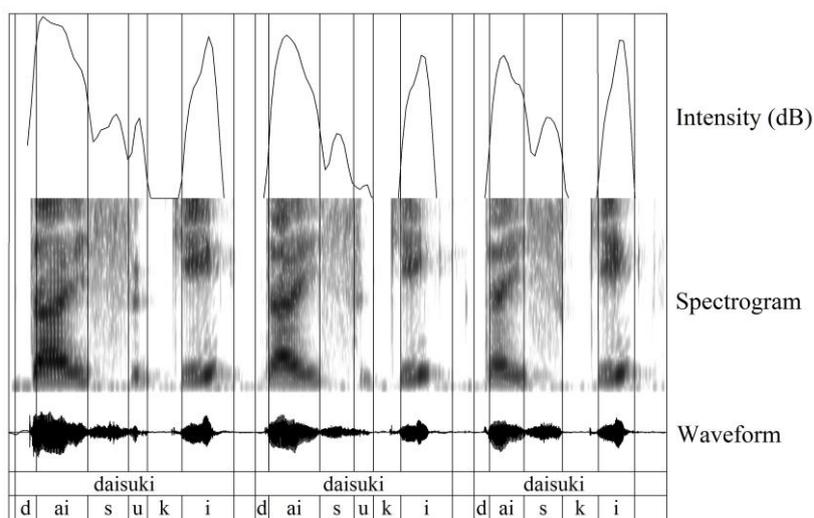


Figure 1: Voiced, devoiced and deleted examples of the /u/ in the Japanese word, daisuki.

## 4. Results

As predicted, participants frequently deleted vowels in the three suffix contexts. This was more frequent when the suffixes were attached to lexical items ( $M = 85\%$ ) than when they were uttered as stand-alone words ( $M = 69\%$ ). Interestingly, both しゅて /ʃute/ [nonce] (19%) and しゅてん /ʃuten/ [red mark] (33%) elicited instances of deletion, creating near homophony between して /ʃite/ [IMP] (83%) and しゅて /ʃute/ [nonce]. As expected, した /ʃida/ [fern] and しで /ʃide/ [paper streamer], which were included to test the influence of having a voiced consonant follow the vowel in question, prompted no instances of deletion. Words with similar but not identical environments to the suffixes did elicit deletion, however this was much more common with medial vowels in すき /suki/ [like] (64%) and だ いすき /daisuki/ [really like] (78%) than in word final vowels in からす /karasu/ [crow] (17%) and まし /maʃi/ [better] (17%). Those words included as instances where non-high vowels undergo devoicing provided few elicitations of this behaviour: ころ /kokoro/ (6%) [heart], ほこり /hokori/ (8%) [pride], and はか /haka/ (0%) [grave].

There was very little difference between the behaviours of those participants from regions where vowel reduction is common and those from regions where it is reportedly infrequent. Participants from the Kansai and Shikoku regions where reduction is infrequent, showed the same level of deletion as those from regions where reduction is common, both groups displayed deletion in 35% of tokens. While these results do show no effect of regional dialect on citation form, one participant exhibited considerable individual difference which warrants consideration. This participant, who indicated that she spent her childhood in Kanagawa just south of Tokyo, produced deletion in only two instances. This individual was the only participant not to delete the /i/ in たべました /tabemajita/ [ate] and はなして /hanajite/ [speak IMP]. After the experimental procedure, the first author conversed with the participant in Japanese, and she was then observed to delete vowels in these contexts in casual conversation suggesting that the participant was hyperarticulating during the experiment.

Finally, a difference was observed between the suffix environments, such that して /ʃite/ underwent deletion most frequently ( $M = 88\%$ ), followed by した /ʃita/ ( $M = 81\%$ ), and finally ます /masu/ ( $M = 78\%$ ). The small data set does not warrant statistical analysis, but we speculate that frequency of occurrence in Japanese is a likely explanation for this pattern. We thus counted the occurrences of each of the three suffixes in a balanced corpus of contemporary written Japanese [9], to examine whether there were differences in the frequency of the three suffixes in Japanese, consistent with the pattern observed in our data. In 105 million words, して /ʃite/ [IMP] occurred most frequently (1,104,381), followed by した /ʃita/ [PST] (860,509), and lastly ます /masu/ [PRS] (664,874). This suggests a relationship between vowel deletion and the frequency at which each suffix occurs, although a larger study is required to test this hypothesis more systematically.

Table 1. Deletion and devoicing results. Words are presented in Hepburn Romanisation and environment indicates the environment of the vowel being monitored.

/i/ Words	Environment	Deletion	Devoicing
shiteru	ʃ_t	89%	--
tabemashita	ʃ_t	92%	--
hanashite	ʃ_t	92%	--
ashita	ʃ_t	83%	--
shita	ʃ_t	67%	--
shite	ʃ_t	83%	--
shida	ʃ_d	--	--
shide	ʃ_d	--	--
mashi	ʃ_#	17%	--
nichi	C_#	25%	--
/u/ Words	Environment	Deletion	Devoicing
shuten	ʃ_t	33%	--
shute	ʃ_t	19%	--
suki	s_k	63%	--
daisuki	s_k	78%	3%
imasu	mas_#	67%	--
tabemasu	mas_#	83%	--
hanashimasu	mas_#	83%	--
u	mas_#	83%	--
karasu	s_#	17%	--
masu	mas_#	56%	--
moku	C_#	--	41%
/o/ Words	Environment	Deletion	Devoicing
kokoro	C_CV	3%	6%
hokori	C_CV	--	8%
/a/ Word	Environment	Deletion	Devoicing
haka	C_CV	--	--
Foils	Environment	Deletion	Devoicing
ima	--	--	--
ichiban	--	--	--
kasa	--	--	--
kaze	--	--	--
mizu	--	--	--
yasha	--	--	--
tusbasa	--	--	--

## 5. Discussion

The results of the present study show that Japanese vowel deletion is a frequently occurring phenomenon, even in words in citation form. Indeed, contrary to research that has found that devoicing does not occur in careful or slow speech, we show that Japanese speakers produce devoiced vowels also in (reasonably carefully monitored) citation form. While devoicing was rare for low vowels (e.g., ころ /kokoro/ [heart] 5.56%), ほこり /hokori/ [pride] (8.33%), and はか /haka/ [grave] (0%)), the only item with a high vowel in a devoicing environment, もく /moku/ [wood], elicited devoicing 41.67% of the time. These findings thus contrast with cross-linguistic data indicating that it is uncommon that speakers elide vowels in careful, planned speech, and the results are particularly interesting because of Japanese phonology's strict

adherence to a syllable structure that does not allow for non-nasal consonant clusters.

Our results also show that deletion patterns in Japanese suffixes may be influenced by a number of factors. Of the suffix environments, *して* /ʃite/ [IMP] and *した* /ʃita/ [PST] elicited deletion more frequently than *ます* /masu/ [PRS]. This might be the result of the /i/ being more susceptible to deletion than the /u/, the palato-alveolar fricative /ʃ/ being more likely to elicit deletion than the alveolar /s/, or the difference between medial and final vowels. However, all three suffixes were more likely to be deleted when attached to existing lexical items, indicating that there may simply be an implicit understanding that some lexical items require deletion or that the morphological status of the morpheme influences deletion. These possibilities are supported by the relatively rare deletion of /u/ in *からず* /karasu/ [crow] (17%), when compared to the *ます* /masu/ [PRS] elicitation (72%). This suggests the possibility that lexical frequency is a factor in the likelihood of deletion: vowels are more likely to be deleted in the three suffix environments because these occur very frequently in Japanese discourse [9].

Additionally, vowel deletion in the minimal pair *して* /ʃite/ [IMP] and *しゅて* /ʃute/ [Nonce] allows for the opportunity to examine the co-articulatory features of each vowel on the preceding fricative. Figure 2 shows the production of these two words by one of our participants. Interestingly, the F2 frequencies appear to be the same, these were measured at around 2200Hz which is to be expected with the high front vowel, /i/, but not the back vowel, /u/. There is, however, some F2 resonance lowering towards the end of the fricative in the *しゅて* /ʃute/ [nonce] example, which is likely the result of protrusion from lip-rounding. It should also be noted that the /ʃ/ in *しゅて* /ʃute/ [nonce] also achieved a higher intensity than the fricative in *して* /ʃite/ [IMP].

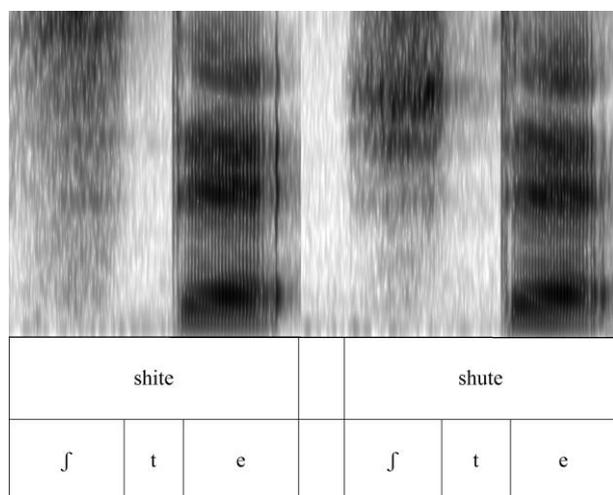


Figure 2: *Minimal Pair, して* /ʃite/ [IMP] and *しゅて* /ʃute/ [Nonce].

Finally, our data provides no support for previously reported regional differences in deletion/devoicing patterns in Japanese. Indeed, we found no significant difference between the behaviour of those from regions of Japan where reduction is common and those who are not. It is likely that this is the result of speakers from the Kansai and Shikoku regions adopting features of the higher prestige Standard Japanese due

to the knowledge that their behaviour was being observed. Standard Japanese, the dialect spoken in centres of commerce and government, most closely resembles dialects spoken in and around Tokyo, a characterising feature of which is vowel reduction.

## 6. Conclusion

Deletion is generally regarded as a feature of fast, casual or connected speech. However, in Japanese, vowel deletion occurs when speakers produce planned, careful utterances. Vowel deletion and vowel devoicing are likely separate, although related, phenomena. Both behaviours occur in similar environments, deletion occurs after fricatives while devoicing occurs after plosives. While further research is required, another point of distinction is likely the relative frequency at which they occur. These behaviours can be distinguished by examining spectrographic evidence. While some may still argue that the vowel is masked by the noise of the preceding fricative, we present figure 1 which clearly shows that a devoiced vowel will result in an increase of frequency following a fricative.

We acknowledge that this is a preliminary report, but argue that frequency at which words occur in Japanese discourse appears to have a direct effect on the deletion of vowels. Those words containing the frequently occurring verb suffixes were more likely to elicit deleted vowels. This can be framed within the principle of least effort, which proposes that languages will move towards requiring less energy from the speaker. However, this tendency has to be balanced against the need for listeners to be able to differentiate between signifiers. The deletion of vowels in each suffix does not lead to homophony because the vowels have low informativeness—they are predictable and have a near-zero entropy. Therefore, deleting these vowels does not result in the loss of information in the signal.

## 7. Acknowledgements

We would like to thank our Japanese participants as well as those who helped to organise participants: Eri Suzuki and Takako Ikeda. We would also like to thank two anonymous reviewers for helpful suggestions and comments.

## 8. References

- [1] Fujimoto, M., “Vowel Devoicing”, in H. Kubozono [Ed], *Handbook of Japanese Phonetics and Phonology*, Walter de Gruyter, 2015.
- [2] Labrone, L., *The Phonology of Japanese*. Oxford University Press, 2012.
- [3] Vance, T. J., *An Introduction to Japanese Phonology*, New York Press, 1987.
- [4] Shibatani, M., *The Languages of Japan*, Cambridge University Press, 1990.
- [5] Sugito, M., *Nihongo Onsei no Kenkyu 3: Nihongo no oto*, Izumi Shoin, 1996.
- [6] Tsujimura, N., *Introduction to Japanese Linguistics*, Wiley Blackwell, 2013.
- [7] Itô, J. and Mester, R. A., “The Phonology of Voicing in Japanese: Theoretical Consequences for Morphological Accessibility”, *Linguistic Inquiry*, 17(1), 49-73, 1986.
- [8] Tsuchida, A., “Fricative-vowel Coarticulation in Japanese devoiced syllables: Acoustic and Perceptual Evidence”, *Working Papers of the Cornell Phonetic Laboratory*, 9, 183-222, 1994.
- [9] National Institute for Japanese Language and Linguistics, *Kotonoha: Balanced Corpus of Contemporary Written Japanese*. Retrieved from [www.kotonoha.gr.jp/shonagon](http://www.kotonoha.gr.jp/shonagon), 2009.