High vowels devoicing and elision in Japanese: a diachronic approach

Francesca Pinto

Università degli studi di Roma "La Sapienza" - Fondazione Ugo Bordoni
francesca.pinto87@gmail.com

ABSTRACT

The present study will focus on occurrence of fully voiced, devoiced and elided high vowels in Japanese. High vowels elision in Japanese is supported by clear acoustic evidences, even though it is not globally accepted by scholars in the field. Elided vowels, hence, are considered here to have their own status.

Moreover, this study is conceived in order to state differences and analogies among four dialects and between five age groups. In fact, the main interest of this research is to prove that high vowel elision may be considered as an innovative phenomenon. Consequently, it is expected that a diatopically and diachronically based study may suggest the drift of Japanese language about the said phenomenon. Therefore, it may be possible to hypothesize changes on the strict syllable structure deriving from Japanese writing system, mainly composed of syllabic morae, which conditions mother tongue perception of Japanese phonology and phonotactics.

Keywords: High vowels devoicing, elision, Japanese dialects, diachrony.

1. INTRODUCTION

Japanese high vowels devoicing is a well known and studied phenomenon, examined by many native and non-native scholars [e.g. 8, 11, 15] since the early stages of literature on Japanese linguistics.

High vowel devoicing is a phenomenon for which vowels are pronounced without the vibration of the vocal folds and where oral organs configuration is more similar to that of whispering. They are hence, voiceless. Vowel devoicing is more likely to occur in specific phonotactic contexts, namely between two voiceless consonants or between a voiceless consonant and a pause. Factors like speech rate may also have a strong influence.

Under the same circumstances high vowel elision may occur as well. Depending on a higher speech rate, on the quality of the consonantal environment and on the word recurrence, this latter phenomenon may have an even higher degree of occurrence as compared to vowel devoicing. In fact, it is quite evident that some specific contexts require vowel elision rather than fully voiced or devoiced vowels. The reason for this distribution is partly due to the ease in pronouncing certain consonantal clusters instead of others – a cluster (resulting from the elision of an interconsonantal vowel) composed of a palatal fricative plus an alveolar occlusive is easier to pronounce than one composed of two occlusives. Other than that, words frequency plays an important role in determining vowel elision regularity: words like *desu* 'to be' or *shita* '(verbal morpheme for past tense)' have a high degree of frequency in Japanese and their realization with high vowel elision has gradually become a stable habit.

Notwithstanding clear and neat phonetic evidences that prove vowel elision, traditionally only vowel devoicing is accepted by scholars in this field, especially by the native ones. It is possible to assume that the reason for that is mainly due to the influence exerted by the Japanese writing system. Apart from Chinese ideograms and Latin characters, Japanese language is conveyed by kana, a syllabic writing system where each symbol stands for a mora composed of a vowel, or a consonant plus a vowel, with the exception of the symbol <n>, the only consonant that can occur alone, but not in word initial position. In other words, Japanese kana has influenced the perception of Japanese phonology and phonotactics so much that even scholars are lead to think that only open syllables are possible – with the said exception – and hardly acknowledge vowel elision and consonantal clusters that derive from this process.

In the present paper it is claimed that vowel elision has its own specific phonological status. Furthermore it can be considered as an innovative feature in Japanese language. The main proof to support this hypothesis relies upon dialect differences. It is well known [5, 11] that Kansai dialects, in the South-centre of Japan, mainly represented by Ōsaka and Kyōto, are highly conservative and tend to keep archaic features, as compared to Standard Japanese. In this area a fully voiced realization of /i, ɯ/ is more frequent, instead of vowel devoicing or elision and it is partly because here rounded variant of the high back vowel is used, whereas in Standard Japanese its unrounded counterpart is used, making the phone more audible and neater. Moreover, younger speakers are more inclined to vowel reduction than adults or elderly.

For this reason, an analysis of this phenomenon occurrence through a diachronic study may lead to a new perspective on this topic and make a

conjecture about Japanese language future development.

In this paper, then, emphasis on diachrony is pursued through both analyses on different age groups and regions. It is expected that a comparison among four diatopic and five diachronic variants may allow hypothesis on two main topics. It would be possible to propose some accurate assumption about a description of dialects distribution from a diachronic point of view. Other than this, it is hoped that analysis from the present paper will help in outlining the possible phonological drift that Japanese is following, at least with regards to high vowels treatment and, eventually, to phonotactic implications.

2. METHODOLOGY

2.1. Corpus

The corpus used for the aim of the present study is *The Speech Corpus and Database of Japanese Dialects* (henceforth SCDJD), supervised by Miyoko Sugito during the 90's [14]. This corpus is composed of interviews on Japanese native speakers divided into five age groups, coming from 14 Japanese prefectures. Audio files have been sampled at 48kHz with 16bit accuracy and are based on interviews dealing with isolated words and short sentences, stories, weather forecast, greetings, questions and answers, short conversations, Japanese syllabaries reading and numerals.

Speech samples from four regions, namely one in the North, Hirosaki, one in the South, Kagoshima, and two in the central part of Honshū island, Nagoya and Ōsaka – the latter being representative of the Kansai dialect – have been used here.

2.2. Speakers

In the SCDJD five age groups are present, namely: elderly, aged over 60 years old; adults, aged 40 to 59; youth, aged 20 to 39; middle school students, aged 13 and 14; elementary school students, aged 10 to 12 years old. Even though a scrupulous diachronic analysis would need a wider time frame, it is possible to state that five generations of speakers may anyhow contribute significantly to a diachrony based study.

For the aim of the present research speech samples of twenty speakers per each region have been used. In particular, speech samples are related to at least two speakers per age group and sex.

A total amount of 80 speakers have been analyzed here.

2.3. Collected data

Starting from the corpus transcription, selection have been made on words representative of five phonotactic contexts, involving high vowels in devoicable position, for instance: (1) the case were /ɯ/ stands between /s/ and a pause; (2) when high vowels are between two voiceless occlusive; (3) between two voiceless fricatives; (4) between a voiceless occlusive and a fricative; (5) between a voiceless fricative and an occlusive.

About 200 words for each of the above mentioned contexts – 966 in total – have been selected, extracted and acoustically analyzed.

It is well known that the definition of the concept of 'word' is quite difficult and it is even harder to define what a word is in Japanese, due to the large varieties of more or less bound affixes, particles and other kind of nearly grammaticalized constructions. Therefore, for the aim of the present study constructions like nouns plus a particle, verb plus inflectional suffixes, adverbs or conjunctions in isolations are considered as single units. Construction where prosodic and syntactic boundaries might not coincide, like nouns followed by a copula, have been excluded by the selection, as well as nouns followed by three or more particles [3]. Moreover, only units with no pitch accent on the studied segments have been chosen.

2.4. Data analysis

The focus of the analysis was to discriminate the quality of the pronounced vowel and to state whether it is fully voiced, devoiced or elided. A voiceless vowel can be easily distinguished from a fully voiced vowel for a series of acoustic peculiarities: it does not display any trace of the voice bar at the bottom of the spectrogram; shows a sensitive increase of both F1 and F2; formants look less neat and sharp; pitch is absent; signal is aperiodic; overall, the phone length is shorter. On the other hand, when a vowel is elided, there is the complete absence of the phone with all its characteristics.

Based on the above mentioned parameters, high vowels in devoicable position of each word have been analyzed and, then, divided into three categories: fully voiced, devoiced and elided, in accordance with their acoustic nature.

Resulting data have been statistically analyzed and subcategorized according to the four chosen Japanese dialects and to the five age groups represented in the corpus.

Acoustic analyses have been conducted using Praat [4].

3. RESULTS

Data results are shown in the following table. Main parameters taken into account to compose this table are diatopic variance, diachronic variance and high vowel articulatory manner. Other parameters like vowel quality, consonant quality surrounding vowels or phonotactic context are not taken into account here.

In this phase, then, results obtained from the above mentioned acoustic analysis are combined and normalized. The displayed percentages are cumulative and subcategorized into groups according to age and dialectal area. To make results more visible, highest percentage within each frame is highlighted in grey, having the meaning of the most preferred articulatory manner among the three contemplated.

Table 1: Analyzed data on the diachronic and diatopic distribution of fully voiced, devoiced and elided vowels.

		Hir.	Nag.	Ōs.	Kag.
Eld.	**Voiced**	31,58	38,25	85,02	18,50
	Devoiced	44,29	22,84	7,27	38,34
	Elided	24,13	39,68	7,27	43,16
Adult	**Voiced**	18,89	21,21	55,96	3,39
	Devoiced	43,88	27,22	13,62	40,73
	Elided	37,23	51,57	30,42	55,87
Youth	**Voiced**	16,67	15,25	38,81	3,13
	Devoiced	40,16	31,60	20,77	43,85
	Elided	43,16	53,15	40,42	53,01
Mid. school	**Voiced**	26,22	11,11	31,97	6,31
	Devoiced	39,29	38,24	19,26	41,02
	Elided	34,48	50,64	48,77	52,67
Elem. school	**Voiced**	26,78	26,18	29,16	5,68
	Devoiced	37,54	33,20	18,75	42,57
	Elided	35,67	40,62	52,09	51,75

Data reported in the above table show a good degree of interest. In fourteen out of the twenty sections obtained from data related to dialects and age groups categories, elided high vowels show highest percentages. Moreover, high rate of vowel elision refers to youth in every region. Being this age group the one where high vowel elision is the most preferred articulator manner shall be considered as a matter of great significance. Speakers aged twenty to forty years old belong to that part of the population that is more detached from local language variety and more inclined to use Standard Japanese. This language variety is, in fact, the most used in universities or in working places: in other words, the variety with which

youth have stronger bounds and deal with more frequently.

Looking at the table from a diatopic point of view it will be clear that Nagoya and Kagoshima dialects are the ones were high vowel elision has a stronger impact, with higher percentages that are related to each age groups. On the other hand, in Kansai area – Ōsaka dialect in the table – fully voiced high vowels are most used in the more adult sections of the population, while younger speakers seems to prefer elided articulation of /i, ɯ/. The said data is quite surprising, considering that expected result from this dialect would be a higher degree of fully voiced vowels across all age groups. Despite influence caused by a different vowel timbre on the degree of elision recurrence, fully voiced high vowels in Ōsaka dialect come at a second place, considering occurrence order.

The dialect where vowel elision is the less preferred articulation manner is the northern one, Hirosaki. In this area, among all the vowels in a weakening favourable position, devoicing of /i, ɯ/ is the phenomenon which is most likely to take place. As mentioned above, a higher percentage of vowel elision is only shown in the youth group age.

In general, looking at the highlighted percentages shown in the table, it is quite evident that, despite being higher, they do not really outclass other articulation manners. Almost all the values, in fact, stand between 37% and 55%; in other words, from one third to about a half of the total amount of analyzed vowels. It may not be surprising that the lowest highlighted percentage is referred to the occurrence of devoiced vowels in the elementary school speaker group in Hirosaki prefecture; also, in this dialect percentages are overall quite homogeneous and not very far apart from each other. In Nagoya and Kagoshima prefecture values are a little higher, surpassing the half of the amount of analyzed vowels in all age groups except elderly and elementary school students in Nagoya prefecture. On the other hand, data referred to Ōsaka are quite interesting in the sense that here data referred to elderly show a very high percentage, about 85%, of fully voiced vowels, contrary to all other categories, where percentages are much more balanced. In the adult group as well, despite being much lower than the elderly group, is anyhow higher than the average.

4. CONCLUSIONS

On the basis of what is exposed above, considering Standard Japanese as a touchstone of dialects here represented and vowel elision as an innovative phenomenon of spoken Japanese, it is possible to outline some tendencies. Varieties in the North-centre – here represented by Nagoya dialect – and in the South – Kagoshima dialect – seem to be less distant to Standard Japanese, where high vowel

elision has quite a high frequency. Ōsaka and Hirosaki dialects come after.

It was already pointed out that difference in the timbre of the back high vowel characterizing Kansai area leads to a higher frequency of fully voiced realization of high vowels, as compared to other dialects. Despite this, acquaintance with such an innovative trait shown by younger speakers represents a matter of great importance.

While analyzing these data it should be not forgotten that there exist factors strictly related to the speaker's age that may possibly have a strong influence on the occurrence of vowel devoicing or elision. Among them, the predominant one is doubtless speech rate. Through data analysis it clearly appeared that age groups with a lower speech rate are those of elderly and of elementary school students. Especially in the latter group, speech was kept under control. On the other hand, as expected, youth had a tendency to a higher speech rate.

This factor may have strongly influenced high vowel elision incidence, for in a very rapid and not so accurate speech this kind of phenomena takes place regularly. Apart from this, it was quite evident that aged speakers were keeping – at times with a pronounced pride – phonetic forms and other kind of phenomena typical of their place of origin. Contrastingly, younger speakers have a stronger habit in using forms more similar to Standard Japanese. Ultimately, students and, in particular, elementary school pupils, seemed to pay a very high attention to the ortophony of any word they had to pronounce; this may be very likely attributable to the influence coming from school education, that leads them to observe more the correct pronunciation of every word, also for what concerns pitch accent.

What is exposed here finds comfort in other studies, dealing with a similar topic. Above all, the research conducted by Byun [5] leads to the same direction. In her paper the author discusses the degree of occurrence of high vowel devoicing in the Keihanshin[1] area of Japan, not taking into account vowel elision. Despite that, her results seems to be quite alike, being devoicing rate of elderly significantly lower than that of younger generations. While her study focuses more on the behaviour of high vowels just in the Kansai dialects and on parameters not examined here, like pitch accent or number of morae in the word, it is interesting to note that data analysis within the same area is quite analogous to the one displayed here. Just like in the present paper, vowel reduction does not affect much pronunciation of aged speakers – devoicing rate being around 36% – while it is much more relevant in adults, 62% and represents the prevalent pronunciation manner among young Kansai speakers, 70%. It should be

kept in mind that percentages reported here refers only to vowel devoicing, for the author does not take into account other kinds of vowel reduction. Considering this, it is possible to state that amount of reduced vowel is fairly the same.

Results shown in the present study may be of some utility, assuming that Japanese language is moving phonologically toward the acceptance and the inclusion of elided vowels in its system. Even though this phenomenon is still not accepted by most of the scholars, elided vowel status is not only proved by clear phonetic evidences: phonological prominence of high vowel elision is clearly confirmed by statistic evidences disclosed in this research. Notwithstanding different degrees of occurrence among age groups and diatopic areas, it is possible to state that high vowels represent in most cases the most preferred pronunciation manner.

In conclusion, the present study can give a contribution to the state of the art on Japanese vowels devoicing and elision and support the change from its strict phonotactic structure to a new one, including a higher number of consonantal clusters through the acknowledgement of high vowel elision.

5. FUTURE WORK

In order to have a more complete outline of Japanese dialects distribution from a diachronic perspective, analysis on a larger scale is necessary. Moreover, analyzing signals from old radio or television programs – e.g. before or during the second world war – may offer the chance to observe last century Standard Japanese and can be very beneficial in understanding the drift of high vowel reduction and, finally, of Japanese phonotactics.

6. ACKNOWLEDGEMENTS

7. REFERENCES

[1] Akamatsu, T. 1997. Japanese Phonetics: Theory and Practice. Munchen: Newcastle.
[2] Amino, K., Arai, T. 2009. Dialectal Characteristics of Ōsaka and Tōkyō Japanese: Analyses of Phonologically Identical Words. *ISCA*, 2303-2306.
[3] Arai, T., Warner, N. 1999. Word Level Timing in Spontaneous Japanese Speech. *Proc. of the XIV Int. Cong. of Ph. Sc.* San Francisco, 1055-1058.
[4] Boersma, P., Weenik, D. 2011. Praat: doing phonetics by computer (Version 5.2). University of Amster-

---

1 Keihanshin area refers to Kyōto, Ōsaka and Kobe.

dam.

[5] Byun, H. 2011. Keihanshin niokeru kyōboin no museika. J. of the Phon. Soc. of Japan 15(2), 23-37.

[6] Dupoux, E. et al. 1999. Epenthetic Vowels in Japanese: a Perceptual Illusion? *J. Exp. Psych: Human Perception and Performance*, 25(6), 1568-1578.

[7] Fujimoto, M. 2004. Boinchō to boin no museika no kankei: Tōkyō hōgen washa to Ōsaka hōgen washa no hikaku. *Stud. in the Jap. Lang.* 55(1), 58-60. Tōkyō: The Soc. of Japanese Linguistics.

[8] Frellesvig, B. 2010. A History of the Japanese Language. Cambridge: University Press.

[9] Kashino, M. 2006. Phonemic Restoration : The Brain Creates Missing Speech Sounds. *Ac. Sci. And Tech.* 27 (6), 318-321. Tōkyō : Acoustical Society of Japan.

[10] Kondo, M. 2005. Syllable Structure and its Acoustic Effects on Vowel in Devoicing Environments. In: Van de Werjer, J., Kanjo, K., Nishida, T. (eds.) *Voicing in Japanese*, 229-246. Berlin, N. Y.: Mouton de Gruyter.

[11] Maekawa, K. 1990. Boin no museika. Nihongo to Nihongo kyōiku. Tōkyō: Meiji shoin.

[12] Nusbaum, H. C., Degroot, J. 1990. The Role of Syllables in Speech Perception. *CLS* 26 (2), 287-317. Chicago: Chicago Linguistic Society.

[13] Shibatani, M. 1990. The Languages of Japan. N. Y.: Cambridge University Press.

[14] Sugito, M. (ed.) 1989-1993. The Speech Corpus and Database of Japanese Dialects. Ministry of Education, Science and Culture of Japan.

[15] Vance, T. J. 2008. The Sounds of Japanese. Cambridge: University Press.