Journal of Phonetics 52 (2015) 70–88

Contents lists available at ScienceDirect

Journal of Phonetics

journal homepage: www.elsevier.com/locate/phonetics

Research Article

Continuous versus categorical aspects of Japanese consecutive devoicing

Kuniko Y. Nielsen *

Oakland University, Linguistics Department, 1025 Human Health Building, Rochester, MI 48309-4401, USA

ARTICLE INFO

Article history:
Received 19 December 2013
Received in revised form
13 May 2015
Accepted 25 May 2015

Keywords:
Japanese vowel devoicing
Consecutive devoicing
Speaker variability

ABSTRACT

The phenomenon of high vowel devoicing is almost obligatory in the Tokyo dialect, except for some environments in which complete devoicing is often blocked. One such case is so called consecutive devoicing, where two or more consecutive vowels are in devoicing environments. Although several accounts of consecutive devoicing have been proposed, its linguistic nature is still being debated. The current study presents a detailed investigation of the nature of consecutive devoicing in Japanese, examining the influence of various factors on its likelihood, speaker variability, and phonetic realization. Twenty-four native speakers of Tokyo Japanese produced 30 words containing consecutive devoicing environments. Mixed-effects modeling revealed that several phonetically and phonologically motivated factors simultaneously contribute to the likelihood of consecutive devoicing and the pattern of partial devoicing. Wide intra- and inter-speaker variability was observed, indicating that realization of consecutive devoicing is not always consistent within a speaker, and that the relative weight of conditioning factors may vary across speakers. Vowel duration in consecutive devoicing environments (which reflects partial devoicing) showed a bimodal distribution, indicating that realization of consecutive devoicing cannot be determined solely by its phonetic environment. Taken together, the results demonstrate both continuous and categorical aspects of consecutive devoicing in Japanese.

1. Introduction

In the standard and many other dialects of Japanese, short high vowels /i/ and /u/ tend to be devoiced or deleted when they occur between two voiceless consonants (Han, 1962; McCawley, 1968; Vance, 1987).

(1)	/hikaku/>[ç̥ikakɯ]/[çkakɯ]	'comparison'
	/kokusai/>[kok̥ɯsai]/[koksai]	'international'
	/kikai/>[k̥ikai]/[kʲkai]	'machine'

High vowel devoicing is almost obligatory in the Tokyo dialect, except in some environments where complete devoicing is often blocked. One such case is so called consecutive devoicing, where two or more consecutive vowels are in devoicing environments. Although several accounts of consecutive devoicing have been proposed (e.g., Kondo, 2005; Tsuchida, 1997; Yoshida, 2004), its linguistic nature, in particular whether it is a phonetically driven continuous process or a phonologically driven categorical process, is still being debated. The current study aims to elucidate the nature of consecutive devoicing, by examining the relative contribution of various factors which influences its likelihood, its phonetic realization, as well as inter-speaker variability.

This introductory section is structured as follows: In 1.1, existing theoretical accounts of Japanese devoicing are discussed, followed by a brief overview of phonetic realization of Japanese vowel devoicing in Section 1.2. Conditioning factors of Japanese vowel devoicing are reviewed in Section 1.3, and Section 1.4 presents an overview of consecutive devoicing and its theoretical accounts, followed by a review of reported individual variability in Section 1.5. Lastly, Section 1.6 summarizes the issues and motivates the current experiment. Sections 2 and 3 present the experimental methods and results, respectively, and Section 4 discusses the theoretical implications of the observed results.

* Tel.: +1 248 370 2175; fax: +1 248 370 3144.
E-mail address: nielsen@oakland.edu

0095-4470/$ - see front matter © 2015 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.wocn.2015.05.003

1.1. Theoretical accounts of Japanese vowel devoicing

The phenomenon of Japanese vowel devoicing was traditionally considered a phonological assimilation of the feature [+/−voice] (e.g., McCawley, 1968). More recently, alternative analyses were proposed by Tsuchida (2001) and Varden (1998) employing Optimality Theory (Prince & Smolensky, 1993) and Feature Geometry (Clements, 1985), respectively, in which Japanese voiceless vowels are specified as [+spread glottis] instead of [−voice]. Despite the difference in featural representations, these accounts consider Japanese vowel devoicing a phonological assimilation process.

Alternatively, Jun and Beckman (1993) proposed an account of Japanese vowel devoicing as a gradient phonetic process involving overlap of glottal gestures. According to this view, Japanese vowel devoicing is the result of extreme overlap and blending of the vowel's voiced glottal gesture by the adjacent consonant's voiceless glottal gesture. Their proposal was based on the following observations: (1) mean duration of devoiced syllables are shorter than that of the same type of syllables with voiced vowels (Beckman, 1982); (2) in their spectral analysis, the initial part of the Japanese word /suki/ 'like' with devoiced /u/ was identical to the /sk/ consonant cluster in the English word *ski*, and furthermore, there was no resemblance between the devoiced /su/ syllable in /suki/ and the quintessential devoiced vowel /h/ in the English word *who.* This gestural overlap view was supported by Imaizumi, Hayashi, and Deguchi (1995), which showed that professional teachers of hearing-impaired children reduced their vowel devoicing in order to improve their listeners' comprehension. Imaizumi et al. (1995) proposed three possible ways in which devoicing can be achieved as a result of gestural overlap: (1) the devoicing gestures of the voiceless obstruents can be shifted toward the vowel due to an increase in speech rate; (2) the size of devoicing gestures can be amplified, or (3) a combination of the two.

More recently, Tsuchida (1997) and Fujimoto, Murano, Niimi, and Kiritani (1998) have proposed that Japanese devoicing is a combination of both phonological and phonetic processes. Tsuchida (1997), by examining the muscular activation for glottal abduction and glottal opening patterns for various voicing patterns of vowels using electromyography (EMG), demonstrated that there are two distinct patterns of glottal gesture in devoicing, and proposed that there are two distinct mechanisms for Japanese devoicing: phonological and phonetic. The phonological devoicing is regular and complete, which includes devoicing of high vowels between voiceless consonants (except between voiceless fricatives and consecutive devoicing). The phonetic devoicing is a result of gestural overlap or undershoot, and thus it is irregular in occurrence and is gradient in the degree depending on phonetic factors. She noted that devoicing of non-high vowels, high vowels between voiceless fricatives, and consecutive devoicing fall in this category. The acoustic analysis presented in Varden (1998) supports Tsuchida's proposal that both phonological devoicing and phonetic loss of voicing are simultaneously at work in Japanese.

Fujimoto (2004), by comparing the devoicing patterns of Tokyo speakers and Osaka speakers using photoelectric glottography (PGG), argued that devoicing in the Tokyo dialect is phonologically driven, where no voicing gesture was observed for the devoiced vowel (the arytenoids in a midsagittal image retracted and the glottis in a cross sectional image continuously opened during the devoiced /CVC/ sequence), and occurred regularly without speaker variation. On the other hand, devoicing in the Osaka dialect is phonetically driven, in which a voicing gesture is observed for the devoiced vowel (and there are two glottal opening peaks for the two consonants surrounding the voicing gesture) even if the voicing is not realized, and the rate of devoicing varied among speakers from close to zero up to 100%.

Examining the data from consecutive devoicing cases, Kondo (2005) argued that Japanese vowel devoicing is part of a vowel weakening process, and that the vowel weakening process may be influenced by Japanese syllable structure. She further proposed that two different mechanisms, namely phonetic and phonological processes, control Japanese vowel devoicing depending on the environment.

Cross-linguistically speaking, devoicing tends to affect high vowels that are in prosodically weak positions and adjacent to voiceless consonants, and is often considered to be a part of the vowel reduction process where vowels are first reduced in duration and centralized in quality, then eventually devoiced. In European Portuguese, devoicing is restricted to high vowels in pre-pausal and unstressed positions (Cruz Ferreira, 1999), and to unstressed high vowels in Modern Greek (Dauer, 1980) and Quebec French (Walker, 1984). Similarly, recent work on French vowel reduction (e.g., Torreira & Ernestus, 2011) has shown that unaccented high vowel devoicing (especially /u/) in the phrase-medial position is common. At the same time, Smith (2003) showed that vowel devoicing in standard French also occurs in the prosodically prominent sentence-final position, indicating that devoicing in French is unlikely to be a pure reduction process.

Although devoicing is more probable in unaccented vowels in Japanese (Kitahara, 2001), previous studies suggest that devoicing is not a canonical reduction process in Japanese, either. First, devoicing does not involve apparent centralization of vowels in Japanese (Kondo, 2005). Second, devoicing is not suppressed in clear/careful speech, which is unlikely if its basis is vowel reduction: Martin, Utsugi, and Mazuka (2014) showed that speakers increased devoicing rate in careful, read speech compared to adult-directed speech, and in Fais, Kajikawa, Amano, and Werker (2010), virtually the same devoicing rate was observed between infant-directed speech and adult-directed speech. Further, Ogasawara and Warner (2009) showed that devoicing might facilitate word processing: in word recognition tasks, listeners performed better when vowels were devoiced in the environment where vowel devoicing was expected. Lastly, a negative correlation between devoicing rate and lexical frequency (i.e., more devoicing among low-frequency words) has been reported in large corpus studies (Kilbourn-Ceron, 2015; Maekawa & Kikuchi, 2005), which also suggests that Japanese vowel devoicing is unlikely to be a reduction process, given that more probable words are reduced more in general (Jurafsky, Bell, Gregory, & Raymond, 2001).