Journal of Phonetics 66 (2018) 100–119

Contents lists available at ScienceDirect

Journal of Phonetics

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

Research Article

The lingual articulation of devoiced /u/ in Tokyo Japanese

Jason A. Shaw [a,*], Shigeto Kawahara [b]

[a] Yale University, New Haven, CT 06520, USA
[b] Keio University, Minato-ku, Tokyo 108-8345, Japan

ARTICLE INFO

Article history:
Received 8 December 2016
Received in revised form 7 September 2017
Accepted 21 September 2017

Keywords:
Japanese
Vowel devoicing
Articulatory phonetics
EMA
Phonetic interpolation
Gestural coordination
CV timing

ABSTRACT

In Tokyo Japanese, /u/ is typically devoiced between two voiceless consonants. Whether the lingual vowel gesture is influenced by devoicing or present at all in devoiced vowels remains an open debate, largely because relevant articulatory data has not been available. We report ElectroMagnetic Articulography (EMA) data that addresses this question. We analyzed both the trajectory of the tongue dorsum across VC_1uC_2V sequences as well as the timing of C_1 and C_2. These analyses provide converging evidence that /u/ in devoicing contexts is optionally targetless—the lingual gesture is either categorically present or absent but seldom reduced. When present, the magnitude of the lingual gesture in devoiced /u/ is comparable to voiced vowel counterparts. Although all speakers produced words with and without a vowel height target for /u/, the frequency of targetlessness varied across speakers and items. The timing between C_1 and C_2, the consonants flanking /u/ was also effected by devoicing but to varying degrees across items. The items with the greatest effect of devoicing on this inter-consonantal interval were also the items with the highest frequency of vowel height targetlessness for devoiced /u/.

1. General background

This paper examines the lingual articulation of devoiced /u/ in Tokyo Japanese. A classic description of the devoicing phenomenon is that high vowels are devoiced between two voiceless consonants and after a voiceless consonant before a pause (Fujimoto, 2015; Kondo, 1997, 2005; Tsuchida, 1997 among many others). This sort of description, high vowel devoicing in a particular context, applies to vowels in numerous other languages including e.g., French (Cedergren & Simoneau, 1985; Smith, 2003), Greek (Dauer, 1980; Eftychiou, 2010) Korean (Jun, Beckman, & Lee, 1998) and Uzbek (Sjoberg, 1963), but Tokyo Japanese is arguably the best studied case of vowel devoicing.

There is a large body of work on this phenomenon in Japanese, covering its phonological conditions (e.g., Kondo, 2005; Tsuchida, 1997), its interaction with other phonological phenomena like pitch accent (e.g., Kuriyagawa & Sawashima, 1989; Maekawa, 1990; Maekawa & Kikuchi, 2005; Vance, 1987) and prosodic structure (Kilbourn-Ceron & Sonderegger, 2017), its acoustic and perceptual characteristics (Beckman & Shoji, 1984; Faber & Vance, 2000; Matsui, 2014; Nielsen, 2015; Sugito & Hirose, 1988), and studies of the vocal folds (Fujimoto, Murano, Niimi, & Kiritani, 2002; Hirose, 1971; Sawashima, 1971; Tsuchida, 1997). Fujimoto (2015) provides a recent, comprehensive overview of this research. While now we have a good understanding of many aspects of high vowel devoicing in Tokyo Japanese, there is little data available on the lingual gestures of high vowels when they are devoiced. The only study that we are aware of is Funatsu and Fujimoto (2011), which used EMMA (ElectroMagnetic Midsagittal Articulography) with concurrent imaging of the vocals fold using nasal endoscopy. They found little difference between devoiced and voiced /i/ in terms of lingual articulation. However, this experiment used only one speaker and one item pair (/kide/ vs. /kite/). The study included four repetitions of each item, and offered no quantitative analyses of the data. Our study is intended to expand on this previous work by reporting more data from more speakers and more extensive quantitative analysis.

Why is it important to study the lingual gestures of devoiced vowels? There are a few lines of motivation behind the current study. First, consider Fig. 1, taken from Fujimoto et al.'s (2002)

* Corresponding author.
E-mail address: jason.shaw@yale.edu (J.A. Shaw).

https://doi.org/10.1016/j.wocn.2017.09.007
0095-4470/

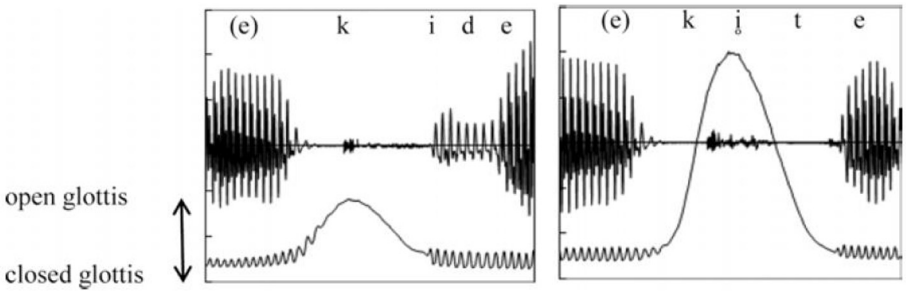

Fig. 1. Degree of glottal abduction for voiced (left) and devoiced (right) vowels in Japanese. The left panel shows a voiceless stop, /k/, followed by a vowel, /i/, and avoiced stop, /d/. For this sequence, there is a single abduction gesture for /k/. The right panel shows a voiceless stop /k/ followed by a devoiced vowel and another voiceless stop, /t/. This sequence also has a single abduction gesture. The magnitude of the abduction gesture in the right panel is larger than twice the size of the abduction gesture in the left panel. Taken from Fujimoto et al. (2002), cited and discussed in Fujimoto (2015).

study, which used nasal endoscopy to image the glottal gestures of high vowel devoicing in Japanese.

Fig. 1 shows that a Japanese devoiced vowel has a single laryngeal gesture of greater magnitude than a single consonant gesture, or even the sum of two voiceless consonant gestures (c.f., Munhall & Lofqvist, 1992 for English which shows the latter pattern).[1] This observation implies that Japanese devoiced vowels involve active laryngeal abduction, not simply overlap of two surrounding gestures. This conclusion in turn implies that Japanese speakers exert *active* laryngeal control over devoiced high vowels (c.f., Jun & Beckman, 1993 for an analysis that relies on passive gestural overlap, to be discussed below). To the extent that Japanese speakers actively control the laryngeal gesture for devoiced vowels, are lingual gestures of devoiced vowels also actively controlled? There are competing views on this matter. On the one hand, it seems logical that active control of a non-contrastive property (allophonic devoicing) would imply active control of a contrastive property (tongue position in the vocal tract). On the other hand, the way that devoicing operates physiologically in Japanese obliterates much of the acoustic signature of lingual articulation. Speakers may not exert active control over aspects of articulation that do not have salient auditory consequences.

The second line of motivation for the current study is the question of whether "devoiced" vowels are simply devoiced or deleted. This issue has been discussed extensively in previous studies of Japanese high vowel devoicing. Kawakami (1977: 24–26) argues that vowels delete in some environments and devoice in others, but he offers no phonological or phonetic evidence. Vance (1987) raised and rejected the hypothesis that high vowels in devoicing contexts are deleted. Kondo (2001) argues that high vowel devoicing is actually deletion based on a phonological consideration. Devoicing in consecutive syllables is often prohibited (although there is much variability: Nielsen, 2015), and Kondo argues that this prohibition stems from a prohibition against complex onset or complex coda (i.e., *CCC). On the other hand, Tsuchida (1997) and Kawahara (2015) argue that bimoraic foot-based truncation (Poser, 1990) counts a voiceless vowel as one mora (e.g., [suto] from [sutoraiki] 'strike', *[stora]).[2] If /u/ was completely deleted losing its mora, the bimoraic truncation should result in *[stora]. Hirayama (2009) makes a similar phonological argument by showing that devoiced vowels' moras are just as relevant for Japanese *haiku* poetry as moras in voiced vowels. However, just because moras for the devoiced vowels remain does not necessarily mean that the vowel is present. The adjacent consonant could conceivably host the mora and syllable—this hypothesis is actually proposed by Matsui (2014), who argues that Japanese has consonantal syllables in this environment (see Dell & Elmedlaoui, 2002 for similar analyses of Tashlhiyt Berber and Moroccan Arabic). Thus, evidence for either deletion or devoicing from a phonological perspective is mixed (see Fujimoto, 2015: 197–198 for other studies addressing this debate).[3]

Previous acoustic studies show that on spectrograms, vowels leave no trace of lingual articulation except for coarticulation on surrounding consonants, which lead them to conclude that vowels are deleted (Beckman, 1982; Beckman & Shoji, 1984; Whang, 2014). An anonymous reviewer questions this finding reported in past work. In principle, a change in the sound source, from modal voicing to turbulence, is independent of the resonance properties of the vocal tract (e.g., Stevens, 1998: 167–168). We might therefore expect to be able to identify formant structure in the aperiodic energy characteristic of devoiced vowels. The reported absence of such structure in past studies may follow from the particular location of turbulent energy sources excited in the vocal tract preceding devoiced vowels in Japanese and their perseverative influence on devoiced vowels. Most of the voiceless consonants preceding devoiced vowels in Japanese are fricatives or affricates that involve turbulence generated by a narrow channel of air in the anterior portion of the vocal tract.[4] The formants produced by these consonants are resonances of the cavity in front of the noise source, i.e., the front cavity (see Stevens, 1998: 176–182 for discussion of the (negligible) effect of the back cavity

[1] Munhall and Lofqvist (1992) investigate the timing and magnitude of laryngeal gestures in the consonants /s/ and /t/ in the sequences *kiss ted* spoken at different speech rates. At slow speech rates two distinct laryngeal gestures can be identified but at faster speech rates the gestures merge into one laryngeal gesture approximating the sum of the two smaller consonantal gestures.

[2] Here and throughout we use the symbol [u] to refer to a broad phonetic transcription. The actual realizations of this vowel in our data in Tokyo Japanese more generally tend not to be as back or as rounded as [u] is strictly defined in the IPA. See Vance (2008: 51) for a detailed description of Japanese /u/. We return to this point when discussing our specific hypotheses below.

[3] Tsuchida (1997) argues that there is "phonological devoicing" as well as "phonetic devoicing".

[4] Japanese lacks singleton /p/, except in some recent loanwords, and /t/ is affricated before high vowels, so the only stop consonant conditioning devoicing that does not involve turbulence generated in the anterior portion of the vocal tract is /k/.