

Inter-Speaker Variations in the Production and Distribution of Devoiced Vowels in Japanese

Yasuko Nagano-Madsen

Japanese Section, Department of Oriental Languages, University of Gothenburg
Västra Hamngatan 3, 41117 Göteborg, Sweden.

Tel: 031-7734844, Fax: 046-104210,

e-mail :Yasuko.Nagano_Madsen@ling.lu.se

1.Introduction

1.1 Vowel devoicing in Japanese

The phenomenon of vowel devoicing has been one of the central issues in phonetic and phonological descriptions of Standard Japanese (Polivanov 1976 [1914], Han 1962, Sawashima 1971, Hirose 1971, Sugito 1988, Maekawa 1989 etc.). At the same time, Japanese has long been nearly an exclusive source of information regarding this phenomenon in linguistic literature. In recent years, however, there has been a growing body of data on vowel devoicing from languages other than Japanese including Montreal French (Cedergren and Simoneau 1985), Korean (Kim 1993, Jun and Beckman 1994), and Turkish (Jennedy 1994), providing an opportunity for cross-linguistic comparison of vowel devoicing. These studies indicate that the phonetic conditions in which vowels are devoiced across several languages are strikingly similar, devoiced vowels are usually high vowels in a voiceless environment.

As far as Japanese is concerned, however, vowel devoicing is primarily a matter of phonology because a devoiced vowel is not a mere by-product of fast or careless speech but is present even in careful slow speech. Sawashima (1971), observing active effort of widening of the glottis for the production of devoiced vowels in Japanese, states that it is not a mere skipping of the phonatory adjustments for the vowel. The fact that devoiced vowels are indicated in pronunciation dictionaries also indicates the status of devoiced vowels within the Japanese phonology.

However, the traditional representation of vowel devoicing as a categorical change from [+voice] to [-voice], has recently been challenged. Sun-Ah and Beckman (1994) argue that devoicing processes such as those found in Japanese and Korean can be better accounted for as a general phonetic process of undershoot, where a vowel's glottal gesture overlaps with a preceding voiceless consonant. They claim that traditional phonological representation cannot explain the typically gradient character of devoicing seen in their data. They consider devoicing as another example of the kind of gradient phonetic effect resulting from more or less subtle adjustments to the magnitude and timing of otherwise invariant

gestural specifications, as in Browman and Goldstein's account of casual speech processes in English (Browman and Goldstein 1990).

One problem in representing vowel devoicing is that it seems to involve much greater variations in its exact phonetic manifestation. One can compare it, for example, with the allophonic variants of /t/ which take the forms [tʃ] and [ts] before /i/ and /u/ respectively, which are much strictly controlled. In addition, realization of vowel devoicing is known to differ both across various dialects of Japanese as well as across speakers within the same dialect (Sugito 1969, 1988, Yoshida and Sagisaka 1990, Yoshida 1992, Nagano-Madsen 1994). Again, no such variations are known for the allophonic variants of /u/.

1.2 *Vowel devoicing factors*

Segmental factors are primary for vowel devoicing in Japanese while other factors such as accent, position in a word etc., which may affect the realization of devoicing, can be considered as secondary. A standard description of vowel devoicing in Japanese is that the high vowels /i/ and /u/ are devoiced between the two voiceless consonants, or between a voiceless consonant and a pause. The description indicates that there are two major requirements for the vowels to be devoiced in Japanese. They are (1) tongue height [+high] of the vowel in question and (2) the lack of periodic glottal vibration [-voiced] in the adjacent consonants. Other factors which have been claimed to influence vowel devoicing include the manner of articulation of prevocalic and postvocalic consonants, accentuation, the presence or absence of another devoiceable vowel in an adjacent syllable, position in the word, utterance size and speech tempo among others.

The influence of accent in the realization of vowel devoicing in Japanese has been known for a long time (cf. Sakurai 1985 for a detailed description). By and large, it has been described that a potentially devoiceable vowel tends to resist devoicing once it is accented. Influence of accent, however, can appear in a different way by shifting accent to either the preceding or the following vowel. Many works using word in isolation form agree that accent indeed has a strong influence in preventing devoicing (cf. Nagano-Madsen 1994 for details).

Only recently have quantitative studies been undertaken into interaction between these factors. Takeda and Kuwabara (1987) followed by Yoshida and Sagisaka (1990), both using ATR word database, have attempted to establish a hierarchy among these factors while Kawai et al. (1993) presented an analysis of the NHK Accent Dictionary. Based on Kawai et al.'s analysis, Nagano-Madsen (1994) suggested that the factor of accent interacts with the manner of the postvocalic consonant. Kondo (1993), in her controlled material, found that the presence or absence of another potentially devoiceable vowel in an adjacent syllable has a critical influence in the realization of vowel devoicing, since devoicing of two successive vowels was not favoured. Some of these findings have opened a new area of close examination in the study of vowel devoicing in Japanese.

1.3 *Inter-speaker differences*

Although devoiced vowels are indicated in dictionaries like the NHK Accent Dictionary, considerable differences have been reported between the professional radio and T.V. announcers. Yoshida and Sagisaka (1990) reported that the total number of devoiced vowels read by three NHK announcers was 538, 601, and 734 respectively in 5449 words. In Nagano-Madsen's analysis of ATR sentence database consisting of 503 sentences, the total number of devoiced vowels varied from 607 to 876 between ten NHK announcers.

Though inter-speaker differences was not the focus of their study, Yoshida and Sagisaka (1990) state that some factors contribute more to the inter-speaker differences. Speakers were found to differ in the distribution of devoiced vowels according to (1) the nature of the prevocalic consonant and (2) whether or not the vowel is word-final. No effects were found from (1) the nature of postvocalic consonant and (2) accent. However, these findings were found not to be applicable to their sentence material in the same study. To date, it is still largely unclear how speakers differ in the production and distribution of devoiced vowels in Japanese.

1.3 *The scope of present study*

Based on the foregoing discussion, the present study examines the inter-speaker variation found in the production and distribution of devoiced vowels in Tokyo Japanese. As a first step, three female speakers, non-professional but teachers of Japanese, age between 48-59, were analysed using controlled corpus consisting of nonsense words embeded in a career sentence. The controlled factors include (1) prevocalic and postvocalic consonant, (2) accent, and (3) presence or absence of another devoiceable vowel in the same word. Prior to the distribution analysis, the phonetic realization of potentially devoiceable vowels were classified into four acoustic prototypes based on the spectrographic patterns. The distribution of acoustic prototypes was then analysed for each speaker.

2. **Material and analysis**

The corpus consisted of nonsense words having three CV syllables, where the segments in the first CVC were varied systematically. All possible combinations of devoiced vowels were included (cf. Table 2). The nonsense words such as *pipaka*, *pitaka*, *pikaka*, were embedded in a carrier sentence "Korewa ____ desu" (this is ____). When the consonant of the second syllable starts with [tʃ, ts, ʃ, ç, f], the following vowel becomes also potentially devoiceable, giving two possibilities for devoiced vowels in succession. Therefore, for those having two potentially devoiceable vowels, another set of words with [ga] was prepared as well (as [pitsuga] in addition to [pitsuka]). Speakers were asked to read these words with two accentual patterns, one having H*L (such as the place name Akita) and the

other LH (such as the place name Mitaka). Speakers exercised few sentences before the recordings.

The entire corpus consisted of 330 words, containing 440 tokens of high vowels in a voiceless environment. The material written in Japanese Katakana was read by three speakers of Tokyo Japanese, all female, aged 48-59 years. All speech material was digitized at 20kHz sampling rates using the Soundscope programme installed on a Macintosh. Acoustic analysis was carried out spectrographically.

3. Results

3.1 Acoustic manifestation

Based on the acoustic analysis of the high vowels produced by the three speakers, four types of acoustic realizations were identified in order to characterize the variations in production. It should be noted that not all of them are clear-cut divisions. It was easy to separate C-type from A-, B-, and D-types by the presence or absence of regular voicing. Difference between A-, B-, and D-types are gradient, varying most consistently in intensity but also in duration and timing in overlapping with the preceding frication phase. A-type has a regular voicing, intensity, and duration while B-type and D-type are reduced successively in that order. A typical sample for each type is shown in Figure 1.

- A. With regular voicing and vowel formant, strong intensity.
- B. Regular voicing is detectable but with greatly reduced intensity, usually also reduced in duration.
- C. Frication phase only.
- D. A further reduced state of type-B, sometimes close to a complete deletion.

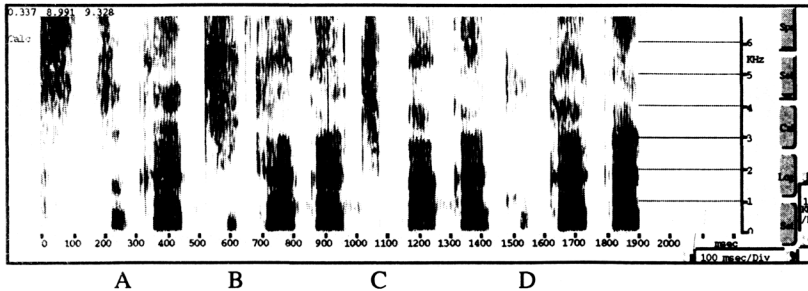


Figure 1. Sample tokens of type-A, B, -C, and -D in the extracted words [sutsuka], [kikaka], [kipaka], and [kukaka] respectively (all with a LHH accentual pattern, speaker S.B.).