The /y/–/w/ Asymmetry in Japanese Loanwords*

Timothy J. Vance and Yuka Matsugu

1. Introduction

Minimal pairs such as *ya* [ja] 矢 'arrow' versus *wa* [ɰa] 輪 'ring' show that modern Tokyo Japanese has two contrasting syllable-initial semivowels. The back semivowel involves the same kind of lip activity as the Japanese high back vowel: compression (not rounding) in careful pronunciation that weakens or disappears in ordinary conversation (Sakuma 1929: 110, Vance 1987: 25–26). We transcribe the two phonemes as /y/ and /w/, reflecting the conventional romanizations.

There is, however, a well-known asymmetry in the distributions of these two phonemes in the modern Tokyo dialect. Confining our attention to syllable-initial CV sequences, there is no question that /ya/, /yo/, /yu/, and /wa/ are allowed and that /yi/ and /wu/ are disallowed.[1] The status of /ye/, /wi/, /we/, and /wo/ is somewhat uncertain, as we will explain below, but in "standard" native and Sino-Japanese vocabulary items, these four sequences do not occur, and according to Bloch (1950: 135), they were absent from what he called the "conservative dialect" of Tokyo Japanese in the 1940s. The asymmetry in Table 1 reflects this description.

Table 1. Syllable-Initial Semivowels in "Conservative" Tokyo Japanese

* /yi/	/yu/	* /wi/	* /wu/
* /ye/	/yo/	* /we/	* /wo/
	/ya/		/wa/

The absence of /yi/ and /wu/ is not at all surprising; CV sequences consisting of a

1 Syllable-initial C/y/V and C/w/V sequences raise additional issues and are beyond the scope of this paper. See Vance (1987: 28–32).

semivowel followed by a homorganic high vowel are universally dispreferred.[2] The same sort of explanation is not available for the other gaps in Table 1, and 8th-century Nara Japanese seems to have allowed all the combinations except /yi/ and /wu/.[3]

Many modern Tokyo speakers seem to have at least some of /ye/, /wi/, /we/, and /wo/ in mimetic items and in certain very formal or very casual expressions; *katakana* spellings, though inconsistent, suggest that these four sequences may occur in recent loanwords (Kawakami 1977: 55, Vance 1987: 26–27). According to Bloch (1950: 160, 162), /wi/, /we/, and /wo/ occurred in loanwords in the "innovating dialect" of Tokyo Japanese in the 1940s, but /ye/ did not. The purpose of the experiment reported in this paper was to begin investigating this issue by determining whether present-day Tokyo speakers actually pronounce some recent loanwords with /ye/ and /we/ rather than with /ie/ and /ue/.

2.   Method
2.1   Materials

We chose four relevant loanwords that are frequently used in Japan today:

(1) a. イェローカード/イェローカード 'yellow card' /yerookaado/ or /ierookaado/

b. ウェブサイト/ウエブサイト 'website' /webusaito/ or /uebusaito/

c. ウェストコースト/ウエストコースト 'West Coast' /wesutokoosuto/ or /uesutokoosuto/

d. ウェルカムボード/ウエルカムボード 'welcome sign' /werukamuboodo/ or /uerukamuboodo/

We matched each of these four items with a name or phrase that unquestionably begins with a VV sequence:

(2) a. 言えるはずない 'cannot possibly say' /ieru hazu nai/
b. 上場鉄造 'Ueba Tetsuzō' /ueba tecuzoo/
c. 上杉謙信 'Uesugi Kenshin' /uesugi keNsiN/
d. 植えるかもしれない 'might plant' /ueru ka mo širenai/

---

2 See Ohala and Kawasaki (1984). This is not to say that such sequences are impossible. English allows them, as in *yield* /yild/ and *wound* /wund/.

3 8th-century Nara Japanese is known as Old Japanese and was represented by *Man'yō-gana* phonograms in the oldest known Japanese documents of any real length. The subsequent history of semivowel+vowel sequences is complicated and beyond the scope of this paper.

As a control pair, we used one item beginning with /yo/ and another beginning with /io/:

(3) a. 寄りかからない 'does not lean against' /yorikakaranai/
b. 庵の向こう 'beyond the hermitage' /iori no mukoo/

For each of these 10 test items (1a–d, 2a–d, 3a,b), we prepared a picture that would elicit it. For example, in the case of the item meaning 'yellow card' (1a), the picture showed a referee with his foot on a soccer ball and holding up a yellow card with his hand. In the case of the two names (2b,c), the illustration was accompanied by a written representation (both *kanji* and *hiragana*) of the given name (*Tetsuzō* or *Kenshin*) but not the surname. In no case was the relevant part of the test item written in *kana* or any form that would directly specify a particular pronunciation, nor did the experimenter (the second author) ever pronounce it.

2.2   Participants

The participants were 13 students from a Japanese university who had just arrived for a short-term course at the University of Arizona's Center for English as a Second Language. They ranged in age from 19 to 21, and all were from Tokyo or the surrounding region. Of these 13 students, only 10 (5 male, 5 female) were able to perform the experimental task without major difficulties, and only their responses were analyzed.

2.3   Task

Each participant was recorded in a sound-treated room. After a brief explanation and some practice, each participant produced a total of 50 tokens in response to the picture prompts described above. Each token consisted of one of the 10 target items in a carrier sentence: /ano hito wa — to iQta yo/ 'That person said — !' The prompts were presented in a different random order for each participant, and each participant produced each of the 10 target items 5 times.

3.   Results
3.1   Duration

We used Praat (the acoustical analysis software package created and distributed at no cost by Paul Boersma and David Weenink) to display a waveform and spectrogram for each token, and we planned to look at three characteristics of each token as potential phonetic correlates of the phonological distinction between semivowel+vowel (SV) sequences and vowel+vowel (VV) sequences. One

characteristic was duration. The VV sequences /ie/, /ue/, and /io/ are disyllabic (and bimoraic) and should therefore have greater duration and more gradual formant movements than the monosyllabic (and monomoraic) SV sequences /ye/, /we/, and /yo/. Each target item in (1)–(3) contains the SV or VV sequence of interest in word-initial position, and the portion of the carrier sentence preceding the target item is the topic phrase /ano hito wa/ 'as for that person'. Consequently, the target item is phrase-initial and likely to be preceded by a pause. In fact, however, in many tokens there is no pause between the /a/ at the end of the topic phrase and the vowel or semivowel at the beginning of the target item. Since there is no reliable way to determine a boundary between the two segments in such tokens, we abandoned the idea of using durational differences to distinguish between SV and VV.

3.2 Glottal Stops

The second characteristic we considered was glottal constriction. A (non-distinctive) glottal stop is likely to precede a phrase-initial vowel, especially following a pause, but a glottal stop is not expected before a phrase-initial semivowel. We sorted our tokens into three categories by inspecting the acoustic displays while listening to the recordings. The tokens in the first category have an unambiguous glottal stop: the acoustic display shows a clear stop burst preceding the formants of the initial segment (vowel or semivowel) in the target item, and we perceive a glottal stop in this position. The tokens in the second category are ambiguous in the sense that the acoustic display shows something other than a clear burst, but we perceive a glottal stop. In almost all of these cases, the acoustic display shows obvious creaky voice. The tokens in the third category lack any trace of a glottal stop: the acoustic display shows what appears to be a smooth onset, and we do not perceive a glottal stop. Figure 1 shows an example in each category.

Category 1
Unambiguous
Glottal Stop
'Uesugi Kenshin'
[speaker 9, token 1]

Category 2
Glottal-Stop-Like
Constriction
'Ueba Tetsuzō'
[speaker 3, token 2]

Category 3
Unambiguous
Smooth Onset
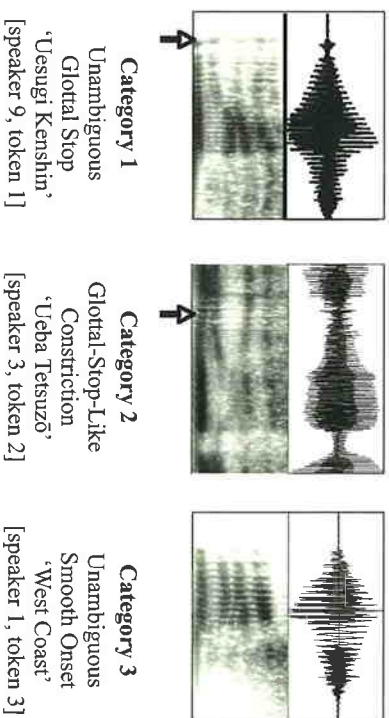'West Coast'
[speaker 1, token 3]

Figure 1. Waveforms and Spectrograms Illustrating the Three Categories of Glottal Constriction

We grouped the tokens into six sets depending on the presumed phonemic form at the onset of the target item:

(4) a. Group 1   /io/   /iori no mukoo/ (3b)
    b. Group 2   /yo/   /yorikakaranai/ (3a)
    c. Group 3   /ie/   /ieru hazu nai/ (2a)
    d. Group 4   /ie/ or /ye/   /yerookaado/ or /ierookaado/ (1a)
    e. Group 5   /ue/   /ueba tecuzoo/ (2b)
                                /uesugi keNšiN/ (2c)
                                /ueru ka mo širenai/ (2d)
    f. Group 6   /ue/ or /we/   /webusaito/ or /uebusaito/ (1b)
                                /wesutokoosuto/ or /uesutokoosuto/ (1c)
                                /werukamuboodo/ or /uerukamuboodo/ (1d)

We then performed three chi-square tests, comparing Group 1 to Group 2, Group 3 to Group 4, and Group 5 to Group 6. The results appear in Table 2, which shows the percentages of each glottal constriction category: unambiguous glottal stop (+), glottal-stop-like constriction (±), and unambiguous smooth onset (−).

Table 2. Glottal Constriction Category Percentages and Chi-Square Results

	−	±	+
GROUP 1 /io/	58	36	6
GROUP 2 /yo/	76	10	14

p < .0005

	−	±	+
GROUP 3 /ie/	52	36	12
GROUP 4 /ie/ or /ye/	58	28	14

n.s. (p = .246)

	−	±	+
GROUP 5 /ue/	49	38	13
GROUP 6 /ue/ or /we/	83	16	1

p < .0005

In terms of the occurrence of glottal stops, the word for 'yellow card' (Group 4) seems to begin with /ie/, since it patterns like the items beginning with uncontroversial VV sequences (Groups 1, 3, and 5). But the words for 'website', 'West Coast', and 'welcome sign' (Group 6) seem to begin with /we/, since they pattern like the item beginning with the uncontroversial SV sequence /yo/ (Group 2).

3.3 Pitch Change

The third characteristic we considered was pitch change in the initial VV or SV portion of the target item in each token. All the target items are either unaccented or accented somewhere other than the initial syllable. A phrase-initial VV in such an item would be described in traditional accounts (Hirayama 1960, NHK 1998, Kindaichi and Akinaga 2001) as beginning with the pitch sequence LH. A phrase-initial SV sequence in a comparable item would be described as having the pitch L, with the following syllable carrying H. Figure 2 illustrates.

```
      L   H             L   H
      |   |             |   |
/u    e  CV.../   /we      CV.../
```

Figure 2. Accent on Phrase-Initial VV and SV Sequences

Using the pitch track generated by Praat, we first located the lowest pitch in the relevant portion of each token and then located the highest pitch after that.[4] We calculated the pitch change (in Hz) by subtracting the lowest pitch from the highest pitch. In one token of /yorikakaranai/ (3a), the pitch was monotonically decreasing throughout /yo/. We assigned this token a pitch change of 0.

We first ran a two-factor within-subjects ANOVA with back versus front vocoid (i.e., /ie/ or /ye/ versus /ue/ or /we/) as one factor, native versus loanword as the other factor, and pitch change as the dependent variable. There was a significant interaction between back/front and native/loanword: $F(1,9) = 16.05$, $p < .005$. We therefore split across back/front and ran two separate one-factor within-subjects ANOVAs with native/loanword as the factor and pitch change as the dependent variable. For the front tokens, which contain either /ie/ or /ye/, the difference between native words and loanwords was not significant: $F(1,9) = 0.54$, $p = .481$. However, for the back tokens, which contain either /ue/ or /we/, the difference

---

[4] In most of the tokens with creaky voice in the relevant VV or SV portion, there were obvious pitch tracking errors. We corrected these by measuring peak to peak in the waveform.

between native words and loanwords was significant: $F(1,9) = 60.25$, $p < .0005$.

In terms of pitch change, the word for 'yellow card' (1a) (mean pitch change = 43 Hz) seems to begin with /ie/, since it patterns like the the item /ieru hazu nai/ (2a) (mean pitch change = 41 Hz), which begins with an uncontroversial VV sequence. In contrast, the words for 'website' (1b), 'West Coast' (1c), and 'welcome sign' (1d) (mean pitch change = 28 Hz) seem to begin with /we/, since they pattern differently from the items /ueba tecuzoo/ (2b), /uesugi keNšiN/ (2c), and /ueru ka mo širenai/ (2d) (mean pitch change = 47 Hz), which also begin with uncontroversial VV sequences.[5]

4. Discussion

Both the glottal-constriction data (Section 3.2) and the pitch-change data (Section 3.3) point to the same conclusion: in recent loanwords in Japanese, the sequence /we/ does occur but the sequence /ye/ does not.[6] Interestingly, the katakana spellings in the authoritative dictionary Kōjien (Shinmura 1998) are consistent with this conclusion. Loan headwords that are common nouns and might be expected to contain /ye/, including the word for 'World Wide Web', appear with the spelling イェ, which would imply /ye/. At the same time, loan headwords that are common nouns and might be expected to contain /we/, including the word for 'yellow card', appear with the spelling ウェ, which implies /we/, and not with the spelling イェ, which would imply /ye/. Remarkably, our results are also consistent with Bloch's (1950: 160, 162) description of the "innovating" Tokyo dialect of the 1940s (see Section 1 above), although we

---

[5] Item (2c) was the only target item accented on the second syllable: /ue'sugi/. We wondered whether the well-known phenomenon of accentual boost (Kubozono 1993: 85–93) would inflate the pitch-change measurements in tokens of this item and yield a misleadingly large mean for the native items beginning with /ue/. To guard against this possibility, we ran our statistical tests again with the tokens of /uesugi keNšiN/ excluded, and the results were virtually identical.

[6] There is, of course, an obvious danger in leaping to the conclusion that our experimental results are representative of modern Tokyo Japanese as a whole, since we collected data on such a small number of target items. Some native speakers of Japanese who attended our public presentations felt that the word for 'yellow card' does begin with /ie/ but that other words, especially proper names such as the Japanese name for the country Yemen, begin with /ye/. These same speakers felt that there are also loanwords that begin with /ue/ rather than with /we/. Only further research on a much larger set of target items can resolve these doubts.

[7] Another authoritative dictionary, Daijirin (Matsumura 1995), uses spellings that imply /ie/ and /ue/ in the same items.

have not provided any evidence for that existence of /wo/ or /wi/.[8]

One could argue, of course, that the real parallel is not between /ye/ and /we/ but between /ye/ and /wo/, both of which involve a semivowel followed a mid vowel of matching backness. It is not easy to find relevant, commonly used loanwords and matched VV items that would allow testing for /wo/ and /wi/, but now that we have a proven technique, we hope to fill these gaps in future work.[9] Our prediction is that speakers like the participants in our experiment will have /wo/ and /wi/ as well as /we/. If so, we will have to face the challenge of explaining why the distributional asymmetry in Table 1 has been only partly eliminated by innovative pronunciations in recent loanwords.

References

Bloch, Bernard. 1950. Studies in colloquial Japanese IV: Phonemics. *Language* 26: 86-125.

Davis, Stuart, and Michael Hammond. 1995. On the status of onglides in American English. *Phonology* 12: 159-182.

Hirayama, Teruo. ed. 1960. *Zenkoku akusento jiten.* Tokyo: Tokyodo Shuppan.

Kawakami, Shin. 1977. *Nihongo onsei gaisetsu.* Tokyo: Ofusha.

Kindaichi, Haruhiko, and Kazue Akinaga. eds. 2001. *Shin meikai Nihongo akusento jiten.* Tokyo: Sanseido.

Kubozono, Haruo. 1993. *The organization of Japanese prosody.* Tokyo: Kurosio.

Matsumura, Akira. ed. 1995. *Daijirin,* 2nd edition. Tokyo: Sanseido.

NHK [Nihon Hoso Bunka Kenkyujo]. ed. 1998. *Nihongo hatsuon akusento jiten,* new edition. Tokyo: Nihon Hoso Shuppan Kyokai.

Ohala, John, and Haruko Kawasaki. 1984. Prosodic phonology and phonetics. *Phonology Yearbook* 1: 113-127.

Sakuma, Kanae. 1929 [1963]. *Nihon onseigaku.* Tokyo: Kazama Shobo.

Shinmura, Izuru. ed. 1998. *Kojien,* 5th edition. Tokyo: Iwanami Shoten.

Tsujimura, Natsuko. 1996. *An introduction to Japanese linguistics.* Oxford: Blackwell.

Vance, Timothy J. 1987. *An introduction to Japanese phonology.* Albany: SUNY Press.

---

8 Davis and Hammond (1995) suggest a similar asymmetry in syllable-initial CGV sequences in American English. They argue that [w] is more consonantal and [j] more vocalic as the G in such sequences.

9 As Stuart Davis suggested to us, it would also be interesting to see how speakers treat relevant loanwords in language games such as the *babibu* game described by Tsujimura (1996: 70-72).