VOWEL DEVOICING IN TOKYO JAPANESE

Mihoko Teshigawara

Department of Linguistics
University of Victoria

1. INTRODUCTION

Japanese is frequently cited as an example of a language with voiceless vowels (Jaeger, 1978; Vance, 1987). First, let us look at examples of vowel devoicing in Japanese, which represent a range of issues addressed in this paper:

(1)	a. /sika/	'deer'	[ʃi̥ka]
	b. /hukahuka/	'soft'	[ɸu̥kaɸu̥ka]
	c. /kasi/	'lyrics'	[kaʃi̥]
	d. /kíka/	'vaporization'	[kí̥ka] or [ki̥ká]
	e. /síhai/	'domination'	[ʃíhai] or [ʃi̥hái]

As a first approximation, we can say that in Japanese the high vowels /i, u/ devoice when they occur between two voiceless consonants ([1a] and [1b]). In addition, high vowels devoice word-finally as in (1c), and we can observe free variation in certain contexts ([1d] and [1e]) where accent (indicated by an acute accent) and vowel devoicing interact. The examples (1a) to (1e) are all from Standard Japanese (henceforth Tokyo Japanese). In this paper, I will attempt to provide a unified phonological analysis of these data.

As a starting point, let us look at how Japanese vowel devoicing has been represented in previous phonological studies. Some early studies (e.g., McCawley, 1968: 127) represented Japanese vowel devoicing by using [–voice] approximately as in (2), and considered it as an assimilation process in the feature [voice]:

(2) $V_{[+high]} \rightarrow$ [–voice] / [–voice] __ [–voice]

There were also some researchers who considered that high vowels to be deleted rather than devoiced (e.g., Ohso, 1973: 13). To my knowledge, no study has investigated phonetic motivations for vowel devoicing.

Recently two major studies were published on Japanese vowel devoicing. Tsuchida (1997) and Varden (1998) investigated vowel devoicing in Tokyo Japanese in depth, using Optimality Theory (Prince and Smolensky, 1993) and Feature Geometry, respectively. What is shared by these two studies is that both Tsuchida (1997) and Varden (1998) assume that Japanese voiceless vowels are specified as [spread glottis][1] (henceforth [s.g.]) instead of [–voice], based on previous researchers' (e.g., Hirose, 1971) and Tsuchida's observations with a fiberscope and electromyography (see Tsuchida [1997] for previous literature using a fiberscope and electromyography). Although these two studies differ in details, generally speaking, we can approximately say that vowel devoicing process is represented as either the

[1] Since Tsuchida (1997) assumes that laryngeal features are privative, her specification of voiceless vowels is [spread glottis], as opposed to Varden's (1998) [+spread glottis].

surfacing of the feature [s.g.] (Tsuchida, 1997) or the spreading of [s.g.] to the voiceless vowel (Varden, 1998).[2] Henceforth, let us call this approach the "spread-glottis approach", in reference to its specification of high vowels for [s.g.]. However, since Japanese lacks a phonological contrast between aspirated and unaspirated consonants, it seems arbitrary to specify voiceless vowels as [s.g.]. Aside from the observation of glottal openings, no phonetic grounding is provided to motivate the [s.g.] specification for voiceless vowels in either Tsuchida (1997) or Varden (1998). In addition, the specification of voiceless vowels as [s.g.] rather than as [–voice] may make it more difficult to generalize the vowel devoicing phenomenon as one of phonological assimilations for pronunciation ease such as sequential voicing and accent shift in compounds. As shown in Teshigawara (2001), there is little motivation for the spread-glottis approach to explaining vowel devoicing. (See Teshigawara [2001] for a detailed discussion of problems with the spread-glottis approach.)

In this paper, an alternative analysis for vowel devoicing in Tokyo Japanese will be proposed by using the feature [–voice] in the framework of Optimality Theory (Prince and Smolensky, 1993). I will draw on Jaeger's (1978) aerodynamic account for high vowel devoicing and other aerodynamic accounts drawn from other researchers' phonetic studies in proposing markedness constraints used in the analysis. In the next section, the basic facts about vowel devoicing in Tokyo Japanese will be introduced. In Section 3, I will propose an alternative phonological analysis for vowel devoicing in Tokyo Japanese, using Optimality Theory (henceforth OT). It will be shown that the aerodynamically motivated constraints can successfully predict correct outputs not only in the canonical devoicing context, but also in word-final position and in the case of accented vowels. Conclusions follow in Section 4.

2. BASIC FACTS ABOUT VOWEL DEVOICING IN TOKYO JAPANESE

Japanese has five vowels, /i, e, a, o, u/, and each of the five vowels has two distinctive lengths, i.e., short and long. No long vowel devoices under any circumstances in any Japanese dialect, which is consistent with Greenberg's (1969) observation that voiceless long vowels are more marked than voiceless short vowels. Among the short vowels, two high vowels /i, u/ are devoiced when preceded and followed by voiceless obstruents, as can be seen in the following examples from Tokyo Japanese.[3]

(3)	a. /sika/	'deer'	[ʃi̥ka]	cf. /siika/	'poetry'	[ʃiika] *[ʃi̥i̥ka]
	b. /kikon/	'married'	[ki̥koɴ]			
	c. /hukahuka/	'soft'	[ɸu̥kaɸu̥ ka]			
	d. /sukii/	'ski'	[su̥kii]			

In each of the four examples (3a) to (3d), the high vowel between two voiceless consonants is devoiced.

In addition to high vowels, K. Sakuma mentions that the non-high vowels, i.e., /e, a, o/ also devoice occasionally, as shown in italics in such words as /h*a*ha/ 'mother', /k*a*karu/ 'to hang' and /k*o*ko/ 'here' (Sakuma, 1929: 231–232, cited in Vance, 1987: 48–49). However, it is also noted that non-high vowel devoicing occurs far less often

---

[2] In Tsuchida's (1997) analysis, all high vowels are specified for [s.g.], and only those that are flanked by two voiceless consonants devoice except for those in some "inhibitory" contexts. In Varden (1998), high vowels flanked by two voiceless consonants receive [+spread glottis] from the preceding consonant.

[3] Throughout this paper, [u] is used for phonetic transcription of /u/ instead of [ɯ]. [ʃ, ʒ] are used to indicate allophones of /s, z/ preceding /i/, and [ç, ɸ] are used for allophones of /h/ preceding /i, u/ respectively. /t, d/ become affricates preceding /i, u/, thus [tʃ, dʒ] appear before an /i/ and [ts, dz] appear before an /u/.

than high vowel devoicing. (See e.g., Venditti and van Santen [1998] for actual devoicing rates of non-high vowels.) Thus, it seems reasonable to say that high vowels devoice in Tokyo Japanese.

Although it has been noted that a high vowel preceded by a voiceless consonant and followed by a voiced consonant can devoice in fast speech (e.g., Beckman, 1994), the devoicing rate in such non-canonical environments is not comparable to that in a canonical environment, i.e., between voiceless consonants. N. Yoshida and Y. Sagisaka (1990, cited in Yoshida, 1998 and 1999) point out that devoiced high vowels preceded by a voiceless consonant and followed by a voiced consonant made up only 4% of devoiced vowels in their data. Thus, it seems legitimate to claim that in Japanese, high vowels devoice between voiceless consonants.

3. ANALYSIS

3.1 Analysis of Basic Facts

Let us analyze the basic facts about vowel devoicing in Tokyo Japanese using OT. To begin with, the fact that high vowels devoice between voiceless consonants is captured by the following context-sensitive markedness constraint:

(4) HVD (HIGH VOWEL DEVOICING) (preliminary version)
$*\mathring{C}\ V_{[+high]}\ \mathring{C}$
No voiced high vowel between voiceless consonants.

This constraint is phonetically grounded. Jaeger (1978) examined the Stanford Phonology Archive, which consists of information on the phonological systems of 221 languages, and found 44 languages with voiceless vowels. Of these 44 languages, 24 devoice only part of their vowel system: of these 24, 20 either devoice only high vowels or preferentially high vowels. Japanese is cited as an example of the latter group. Jaeger observed that the tendency to devoice high vowels is aerodynamically grounded. The relatively narrow oral cavity necessary to produce high vowels (compared to non-high vowels) produces a high supraglottal air pressure. When the supraglottal air pressure becomes too high, the vocal fold closure, which is essential for vocal fold vibration (i.e., voicing), cannot be sustained; therefore, the vocal folds open up, and voicing stops. In addition, the following cross-linguistic perceptual evidence may suggest that acoustic evidence of the influence of vowels on preceding consonants is greater when the vowel is high rather than non-high, thus making voiceless high vowels auditorily less marked than voiceless non-high vowels. In Woleaian, a language where five short vowels (except for the low vowel [a]) devoice before pauses, it is noted that voiceless high vowels, i.e., [i̥, ʉ̥, u̥] are easier to auditorily differentiate than voiceless non-high vowels (Sohn, 1975: 20). Thus, we may assume that it is easier for listeners to retrieve acoustic cues for high devoiced vowels from preceding consonants than for non-high devoiced vowels.

However, although high vowels may devoice in Tokyo Japanese, voiceless vowels are universally more marked than voiced vowels. It has been revealed that cross-linguistically, there is no language with a phonemic contrast between voiced and voiceless vowels (Greenberg, 1969). The marked status of voiceless vowels is captured with the context-free markedness constraint in (5):

(5) NO VOICELESS VOWEL
$*\mathring{V}$
Vowels must not be voiceless.

In addition, in order to prevent unnecessary vowel devoicing, a faithfulness constraint concerning the specification of voice is also required, as in (6):

(6) IDENT-IO (voice)
Correspondent segments in input and output have identical values for [voice].

For allophonic variation, the ranking of the constraints is as follows: the context-sensitive markedness constraint, i.e., (4) HVD dominates the context-free constraint, (5) *V̥, followed by the faithfulness constraint, (6) IDENT-IO (voice), as shown in (7).

(7) HVD >> *V̥ >> IDENT-IO (voice)

The correctness of this constraint ranking is illustrated in tableaux (8) to (12). First, let us consider the case where a voiced vowel is in the input, but a voiceless vowel appears in the output as in (12).

(8) /sika/ 'deer'

Input: /sika/	HVD	*V̥	IDENT-IO (voice)
a. ☞ ʃi̥ka		*	*
b. ʃika	*!		

The candidate (8b), which does not have devoicing on the high vowel /i/, loses to the actual output (8a), since it violates the highest-ranked context-sensitive markedness constraint, HVD. The selected candidate, (8a), violates two lower-ranked constraints, i.e., *V̥ (context-free markedness constraint) and IDENT-IO (voice) (faithfulness constraint). However, this does not affect the outcome since this candidate satisfies HVD, the most highly ranked constraint of the three.

This result should be obtained regardless of different assumptions about the voicing of vowels in the input in order to maintain Richness of the Base, a concept that guarantees that evaluation is performed on a set of candidate outputs, not on the input level, and that no constraints can be stated at the level of input (Prince and Smolensky, 1993). Indeed, the same candidate [ʃi̥ka] is selected when the input contains a voiceless vowel, i.e., /si̥ka/ as in (9). Again it is HVD that determines the outcome, without interference of the lower-ranked constraint, *V̥.

(9) /sika/ 'deer'

Input: /si̥ka/	HVD	*V̥	IDENT-IO (voice)
a. ☞ ʃi̥ka		*	
b. ʃika	*!		

In order to account for the complementary distribution of voiced and voiceless vowels, we should also be able to prove that voicing of the vowels in the inputs does not affect the outcome when there is no devoicing environment. Let us look at two tableaux for the word, /zikan/ 'time,' one with a voiced vowel as its input (10), and the other with a voiceless vowel (11).

(10) /zikan/ 'time'

Input: /zikan/	HVD	*V̥	IDENT-IO (voice)
a. ☞ ʒikan			
b. ʒi̥kan		*!	*

(11) /zikan/ 'time'

Input: /zi̥kan/	HVD	*V̥	IDENT-IO (voice)
a. ☞ ʒikan			*
b. ʒi̥kan		*!	

In both cases, because the context is not relevant to HVD, that is, the high vowel is between a voiced consonant and voiceless consonant, the decision falls to the lower-ranked context-free markedness constraint. As shown in both (10) and (11), the context-free markedness constraint *V̥ is the sole determinant of the output regardless of the voicing of /i/ in the input. Thus, we can conclude that the present ranking is consistent with the concept of Richness of the Base. (Henceforth, only inputs with voiced vowels will be given since Richness of the Base is guaranteed.)

This constraint ranking also predicts correct outputs when a non-high vowel appears between two voiceless consonants as in (12).

(12) /saka/ 'slope'

Input: /saka/	HVD	*V̥	IDENT-IO (voice)
a. ☞ saka			
b. sḁka		*!	*

As is the case with (10) and (11), the context for the application of HVD does not obtain here, therefore the decision falls to the lower-ranked context-free markedness constraint. The candidate (12b), which has a voiceless non-high vowel, is eliminated because of the violation of *V̥. Gratuitous voiceless vowels are not permitted.

In the next three subsections (3.2.1 to 3.2.3), we will see that the present aerodynamic approach to vowel devoicing allows us to provide a coherent approach to some other issues. First, another context for vowel devoicing, where silence follows a devoiceable vowel, i.e., so-called "word-final devoicing", will be analyzed. Then, the relationship between accent and vowel devoicing will be analyzed by using aerodynamically motivated constraints. Lastly, an aerodynamic explanation for the fact that long vowels do not devoice in Japanese will be proposed (3.2.3).

3.2 Detailed Facts about High Vowel Devoicing in Tokyo Japanese

3.2.1 Word-Final Devoicing

In addition to the canonical devoicing context discussed above, there is another context where high vowels devoice. It is generally mentioned that a high vowel preceded by a voiceless consonant and followed by a pause devoices when it has a low tone (e.g., Nihon Hoso Kyokai [henceforth NHK], 1966). For example, /kasi/ 'lyrics' is pronounced as [kaʃi̥] when followed by a pause. However, when followed by another word such as a particle, the voicing of the word-final high vowel depends on the initial consonant of the following word; the /i/ in /kasi/ is devoiced if it is followed by a word starting with a voiceless consonant, e.g., /kara/ 'from', i.e., [kaʃi̥ kara], while it is voiced when followed by a word starting with a voiced consonant, e.g., /demo/ 'even', i.e., [kaʃi demo] (Maekawa, 1989). Thus, in reference to these situations, we can say that a word-final devoiceable vowel (i.e., a high vowel preceded by a voiceless consonant) devoices only utterance-finally; not every word-final devoiceable vowel devoices.

This fact can also be captured by the aerodynamic account of vowel devoicing mentioned earlier. A pause, which is a period of silence, i.e., lack of vocal fold vibration ([–voice]), can be considered as the same as a voiceless consonant. Therefore, we can say that the environment of a preceding voiceless consonant and following pause provides high vowels with the same environment for devoicing as that between two voiceless consonants. In order to allow for the devoicing of a high vowel preceded by a voiceless consonant and followed by a pause, the HVD constraint proposed in (4) is modified as follows.

(13) HVD (HIGH VOWEL DEVOICING) (final version)

$$*\mathring{C}\, V_{[+high]} \left\{ \begin{array}{c} \mathring{C} \\ \text{pause} \end{array} \right\}$$

No voiced high vowel between voiceless consonants or when preceded by a voiceless consonant and followed by a pause, i.e., between a preceding voiceless consonant and a following voiceless period.

The final version (13) takes the place of the preliminary version (4) in the ranking proposed in (7). Now let us look at an example with a final syllable consisting of a voiceless consonant and high vowel.

(14) /kasi/ 'lyrics'

Input: /kasi/	HVD	*V̥	IDENT-IO (voice)
a. ☞ kaʃi̥		*	*
b. kaʃi	*!		

Candidate (14b) is ruled out because of the violation of HVD, which now also prohibits a high vowel preceded by a voiceless consonant and followed by a pause, and the correct candidate (14a) with a voiceless final vowel is selected.

3.2.2 Accent and Vowel Devoicing

Maekawa (1989) mentions both synchronic and diachronic connections between vowel devoicing and accent, which have been observed by previous researchers: the synchronic connection is that accented vowels do not devoice as often as unaccented vowels; and the diachronic connection is that the existence of vowel devoicing caused accent shift in some dialects in Japanese (e.g., Nitta, 1985). Since the diachronic relationship between vowel devoicing and accent is beyond the scope of this paper, it will not be discussed any further in this paper; however, it would be a very interesting topic to pursue in a future study.

First, some basic facts about accent in Tokyo Japanese are introduced. (See e.g., Uwano [1989] and Vance [1987] for more information on accent in the Japanese language.) In Tokyo Japanese, a word can be either accented or unaccented. Each accented word has accent, which is characterized by a pitch fall from high to low. In the following discussion, the last high-pitched mora in an accented word is called the accented mora, indicated by an acute accent mark over the vowel. In Tokyo Japanese, specifying the accented mora in a word is enough to predict the tonal pattern of the rest of the word; if the initial mora is not accented, it receives a low tone, and the succeeding moras up to the accented mora receive a high tone. In the case of unaccented words, there is no such fall in pitch, and the melody starts with a low tone and the remaining moras receive a high tone. The difference between a word with final accent and an unaccented word is not clear when pronounced in isolation, but it becomes clear when followed by another word such as a postposition. For example, a final-accented word /otokó/ 'man' and unaccented word /sakana/ 'fish' have the same pitch pattern LHH when pronounced in isolation, but the difference emerges when followed by a postposition, e.g., /wa/ (topic marker), i.e., /otokó-wa/ LHHL vs. /sakana-wa/ LHHH. For *n*-mora words there are *n+1* accent patterns in Tokyo Japanese. (This number can be correctly predicted by the Prosodic Faithfulness constraints introduced in [16]; see Footnote 5.)

Previous researchers have noted that devoicing of accented vowels tends to be avoided (e.g., Han, 1962; Vance, 1987). However, devoicing of accented vowels has recently become more acceptable in Tokyo Japanese, especially among younger speakers (Tsuchida, 1997). When a word has initial accent and the initial vowel is devoiceable, i.e., a high vowel between two voiceless consonants, there are often two possible pronunciations given to the word, as seen in both NHK (1966) and Hirayama (1960): in one pronunciation the initial vowel is devoiced and accented; in the other pronunciation the initial vowel is devoiced and accent shift or deaccentuation occurs in order to avoid a voiceless accented vowel. Of the two possible data sources I am able to consult, Hirayama's (1960) data will be used in the following analysis, since the pronunciations in Hirayama seem to be closer to the pronunciation of average speakers, although there seem to be some irregularities as well.[4] The following are examples of words that have more than one entry in Hirayama's (1960) dictionary. (The actual pitch patterns for the two pronunciations in each word in [15] are given in parentheses following each pronunciation.)

(15)	a. /kíka/	'vaporization'	[kí̥ka]	(HL)	or	[ki̥ká]	(LH)
	b. /tíketto/	'ticket'	[tʃí̥ketto]	(HLLL)	or	[tʃi̥kétto]	(LHLL)
	c. /sísja/	'branch office'	[ʃí̥ʃa]	(HL)	or	[ʃi̥ʃa]	(LH)
	d. /síseki/	'historical site'	[ʃí̥seki̥]	(HLL)	or	[ʃi̥seki]	(LHH)

In all the examples in (15), the only devoiceable vowel, which is also in the initial syllable of each word, devoices whether it is accented or not. In the first variant, the high vowel is devoiced and accented at the same time. In the second variant, the accent shifts to the following mora as in (15a) and (15b); in (15c) and (15d) deaccentuation occurs and the second variants become unaccented. In either case, however, the voiceless vowel of the second variant is no longer accented and has a low tone.

This is the case of free variation where a single input is mapped onto two grammatical outputs. In order to predict both correct outputs in an OT analysis, we will draw on a concept called "free ranking" (Anttila, 1995; Kager, 1999: 404 – 407), instead of a single deterministic ranking, in which each input is mapped to only one output. Free ranking assumes that two constraints C_1 and C_2 are freely ranked where the evaluation procedure branches: in one branch, C_1 is ranked above C_2; in the other branch the ranking is reversed. In addition to free ranking, we need to propose a set of faithfulness constraints to prohibit accent shift and deaccentuation, which are adopted from Alderete (1999) as in (16), and a context-free markedness constraint to prohibit having voiceless accented vowels as in (17). First, the three Prosodic Faithfulness constraints proposed by Alderete (1999) are introduced.

[4] NHK (1966) is highly prescriptive; for example, almost no voiceless vowels between two fricatives are described, which does not correspond to the situations observed in previous studies such as Hirose (1971) where 70% of high vowels between two fricatives were devoiced.

(16) Prosodic Faithfulness (PROS-FAITH) (Alderete, 1999: 18–19)

a. MAX-PROM: For x a prominence, $\forall x\ \exists x' [\ x \in S_1 \rightarrow x' \in S_2\ \&\ xRx'\ ]$
'Every prominence in S_1 must have a correspondent in S_2.'

b. DEP-PROM: For x a prominence, $\forall x\ \exists x' [\ x \in S_2 \rightarrow x' \in S_1\ \&\ xRx'\ ]$
'Every prominence in S_2 must have a correspondent in S_1.'

c. NO-FLOP-PROM:
For x a prominence, y a sponsor, and z an autosegmental link,
$\forall x\ \forall y\ \forall z$ [x and y are associated via z in $S_1 \rightarrow$
$\exists x'\ \exists y'\ \exists z'$ such that (x, y, z)R(x', y', z') and x' and y' are associated via z' in S_2.
'Corresponding prominences must have corresponding sponsors and links.'

MAX-PROM and DEP-PROM maintain the contrast between accented and unaccented words by prohibiting the deletion of an accent in the input (MAX-PROM), and the insertion of an accent that has no correspondent in the input (DEP-PROM). NO-FLOP-PROM requires that the position of the prominence stay the same in the mapping from one structure to another. Alderete (1999) assumes that in the Japanese grammar, these three faithfulness constraints are ranked in the same position with respect to each other, together constituting the constraint PROS-FAITH, and are ranked higher than alignment constraints that assert a fixed position for prominence structures (e.g., the right edge of the word).[5] However, in the following analysis, we will see that the three constraints are not always ranked in the same position with respect to each other.

Next, let us turn to the context-free markedness constraint that prohibits voiceless accented vowels:

(17) NO VOICELESS ACCENTED VOWEL
$*\acute{\underset{\circ}{V}}$
Accented vowels must not be voiceless.

This constraint is motivated by various factors. High-pitched vowels are produced with greater subglottal pressure than low-pitched vowels (Titze, 1992, cited in Shadle, 1997: 51); thus, from an aerodynamic point of view, we can assume that the greater subglottal pressure of high-pitched vowels prevents them from devoicing. Accented vowels are high-pitched, therefore, they are less likely to devoice than low-pitched vowels. From the viewpoint of laryngeal articulation, Sugito (1998) observed that the glottis adductor muscle was activated during accented syllables, which conflicts with what is necessary for vowel devoicing, i.e., glottal abduction. According to Sugito (1997, 1998), voiceless accented vowels have no pitch, thus no tone realization is possible on the voiceless vowels themselves, and it is the following vowels that realize a steep falling tone, which serves to show that the immediately preceding vowel has accent. Thus, we may say that voiceless accented vowels are acoustically more marked as well.

Since this is a case of free variation, separate constraint rankings are proposed for each of the two variants, i.e., the first containing a voiceless accented vowel and the second manifesting accent shift/deaccentuation. The relevant

---

5 That NO-FLOP-PROM is ranked higher than those alignment constraints ensures that a word with n-numbered moras has n number of accentual contrasts because the accent position in the input must be maintained. As mentioned above, since MAX-PROM and DEP-PROM bring about additional contrast, i.e., the presence or absence of accent, these constraints together yield $n+1$ accent contrasts for n-mora words (Alderete [1999]).

constraints here are (14) HVD, (16) Prosodic Faithfulness constraints, and (17) *V̥. First, let us examine how these constraints are ranked for the words that have two variant pronunciations, the second manifesting accent shift ([15a] and [15b]). In order to allow a voiceless accented vowel to occur in the first variant, *V̥, which is violated by the output form, must be ranked lowest. Prosodic Faithfulness constraints and HVD are equally ranked for the first variant, since we do not have any direct evidence to suggest that they are ordered with respect to one another (see [18]). In the second variant in these words, NO-FLOP-PROM is violated; thus this constraint must be ranked lower than the remaining relevant constraints here, i.e., HVD, *V̥, and the other two Prosodic Faithfulness constraints (MAX-PROM and DEP-PROM), as in (19). Rankings (18) and (19) predict a pair of variants that alternate between a pronunciation with a voiceless accented vowel and one with vowel devoicing and accent shift (i.e., [15a] and [15b]); the constraints that change positions in the two rankings are NO-FLOP-PROM and *V̥.

(18) HVD, MAX-PROM, DEP-PROM, NO-FLOP-PROM >> *V̥

(19) HVD, MAX-PROM, DEP-PROM, *V̥ >> NO-FLOP-PROM

Below, it is shown that rankings (18) and (19) can predict correct outputs for the word /kíka/ ('vaporization'), which has two variant pronunciations, the second manifesting accent shift.

(20) /kíka/ 'vaporization'

Input: /kíka/	HVD	MAX-PROM	DEP-PROM	NO-FLOP-PROM	*V̥
a. ☞ kí̥ka					*
b. ki̥ká				*!	
c. kíka	*!				
	HVD	MAX-PROM	DEP-PROM	*V̥	NO-FLOP-PROM
a'. kí̥ka				*!	
b'. ☞ ki̥ká					*
c'. kíka	*!				

In the upper and lower parts of (20), the optimal candidates are different, as predicted by the two different constraint rankings illustrated therein. In the upper part, (20b) is ruled out since it violates the highly ranked constraint NO-FLOP-PROM by shifting the accent to the following mora, whereas in the lower part, where this same constraint is ranked lowest, (20b') is selected. In both rankings, neither (20c) nor (20c') is selected, because they violate the highly ranked constraint HVD.

In order to predict correct outputs for the words that have a second variant with deaccentuation ([15c] and [15d]), we need to propose another constraint ranking. Since MAX-PROM is violated in the second variant of these words, this constraint must be ranked lower than the rest of the relevant constraints (i.e., HVD, *V̥, DEP-PROM, and NO-FLOP-PROM), as in (21).

(21) HVD, DEP-PROM, NO-FLOP-PROM, *V̥ >> MAX-PROM

Constraint rankings (18) and (21) account for the variant pair that alternates between a pronunciation with a voiceless accented vowel and one with vowel devoicing and deaccentuation (i.e., [15c] and [15d]); the constraints that switch positions in these two rankings are MAX-PROM and *V̥. (22) shows that rankings (18) and (21) can predict correct outputs for words that have a second variant with deaccentuation, such as /sísja/ ('branch office').

(22) /sísja/ 'branch office'

Input: /sísja/	HVD	MAX-PROM	DEP-PROM	NO-FLOP-PROM	*V̥
a. ☞ ʃi̥ʃa					*
b. ʃi̥ʃa		*!			
c. ʃiʃa	*!				
	HVD	DEP-PROM	NO-FLOP-PROM	*V̥	MAX-PROM
a'. ʃi̥ʃa				*!	
b'. ☞ ʃi̥ʃa					*
c'. ʃiʃa	*!				

Here again, (22b) is ruled out in the upper part of the tableau, because the deletion of the accent violates MAX-PROM, whereas (22b') is selected in the lower part, because it satisfies one of the most highly ranked constraints, *V̥.

Further support for the present aerodynamic approach comes from examining specific consonantal contexts. Depending on the consonants preceding and following the initial accented devoiceable vowel, vowel devoicing and accent may depart from the patterns discussed thus far in this paper: different devoicing patterns are observed in other consonantal environments. A survey was conducted for this paper using Hirayama's (1960) dictionary (consisting of approximately 100,000 words) in order to examine the relationship between consonant environments and vowel devoicing patterns. Previous studies such as Tsuchida (1997) suggest that high vowels between two voiceless fricatives and those followed by an allophone of /h/ are less likely to devoice than those between two plosives. Thus, the objects of the survey were limited to words beginning with the following four types of sequences containing $\c{C}_1$ $_{[+high]}$ $\c{C}_{2,}$ as in Table 1 (23):

(23)

	$\c{C}_1$	$\c{C}_2$
a.	plosives	/s/ or /sj/
b.	plosives	/h/
c.	continuants	/s/ or /sj/
d.	continuants	/h/

Table 1: Objects of the survey: words beginning with the four types of sequences containing $\c{C}_1 V_{[+high]} \c{C}_2$

Only words that have a second vowel that is not devoiceable, i.e., a non-high vowel, a long vowel, or a vowel followed by a voiced consonant were examined; consecutive devoiceable environments were excluded from this survey. Table 2 in (24) shows the percentages of words containing a voiceless accented vowel compared to words with a devoiceable accented vowel in the initial mora. In other words, Table 2 shows the percentages of words that have the same vowel devoicing patterns as we have seen in the previous discussion from (15); there are two variants: one devoices the accented devoiceable vowel without any accent shift, while the other devoices the devoiceable vowel with accent shift or deaccentuation.[6]

[6] One of the samples that has a sequence of "continuant – high vowel – /s/" has different vowel devoicing patterns:
(i) /sísa / 'suggestion' [ʃisa] or [ʃi̥sa]
This is included in the category that allows devoicing of the accented vowel.

(24)

C̥$_2$ \ C̥$_1$	/s/ or /sj/	/h/
Plosives	100.0 (7/7)	100.0 (3/3)
Continuants	71.4 (10/14)	14.3 (1/7)
Sum	81.0 (17/21)	40.0 (4/10)

Table 2. Percentages and numbers of words that have a voiceless accented vowel compared to those that have a devoiceable accented vowel in the initial mora. The figures in parentheses show the actual numbers of words in each category.

Despite the observations made in previous studies such as Tsuchida's (1997),[7] which mention that vowel devoicing between fricatives and before /h/ are equally prohibited, the results seem to suggest that these two environments are different. While devoicing high vowels between a continuant and /s/ or /sj/ is still common, devoicing high vowels between a continuant and /h/ is far less common. Thus in the following discussion, only vowels before /h/ are assumed to be undevoiceable, and the four examples that allow devoicing of the accented vowel are excluded from the analysis.

The pronunciations for those words that do not allow the devoicing of accented high vowels before /h/ are as follows:

(25) a. /sîhai/ 'domination' [ʃihai] or [ʃi̥hái] b. /sîheɴ/ 'poetry' [ʃiheɴ] or [ʃi̥heɴ]

Unlike in (15), the first variant of each word in (25) does not devoice the initial accented vowel. In the other pronunciation, however, the pattern observed is the same as in (15); the initial vowel is devoiced with accent shift (25a) or is devoiced and deaccentuated (25b).[8] In order to predict the correct outputs, we need to add a constraint to prohibit the occurrence of a voiceless accented vowel before /h/ or its allophones.

(26) *V̥́C$_{[+cont, -strid]}$
No voiceless accented vowels may precede [h, ç, ɸ].[9]

This constraint is phonetically grounded. [h] takes much more airflow to produce compared to other voiceless fricatives. According to Shadle (1997: 44), the volume flow rate for [h] may be 1,000 to 1,200 cm^3/s compared to a rate of 200 to 400 cm^3/s for typical voiceless fricatives. We may assume that the greater airflow necessary to produce [h] would increase the airflow during the production of the preceding vowel anticipatorily; the increased airflow would

---

[7] Tsuchida (1997) used a revised version of NHK (1966) (NHK, 1985) as a data source. While I am not able to consult that particular version of the dictionary, from Tsuchida's (1997) analysis, it is clear that NHK's dictionary almost exclusively bans the devoicing of high vowels between two fricatives and before an allophone of /h/, whether they are accented or not.

[8] The same patterns are observed for a minority of words consisting of the sequence "continuant – high vowel – /s/ or /sj/":

(i) a. /sísjoo/ 'teacher' [ʃiʃoo] or [ʃi̥ʃóo] b. /sjúsa/ 'chief examiner' [ʃúsa] or [ʃu̥sá]

[9] As mentioned in Footnote 3, /h/ becomes [ç, ɸ] when it precedes /i, u/ respectively. However, only two out of ten samples in the data in Table 2 have an /u/ following /h/, and none has an /i/ following /h/. Thus, we may say that the constraint in (26) mostly deals with [h] rather than [ç, ɸ].

lead to voice the preceding vowel. While /h/ appears as [ç, ɸ] before /i, u/ respectively, possibly resulting in a lower volume flow rate, [ç, ɸ] may also appear as [h] (Tsuchida, 1997; Vance, 1987); and as mentioned in Footnote 9, there are fewer environments in which [ç, ɸ] appear, compared to environments in which [h] appears. In addition, as mentioned earlier, it is the following vowel that realizes the accent with a steep falling tone when the initial accented vowel is devoiced. Thus the sequence of "V̥/h/" followed by a non-devoiceable vowel with a falling tone, i.e., [V̥h, V̥ç, V̥ɸ], would require more articulatory effort than, say a sequence of "V̥/h/" followed by a level tone, i.e., unaccented [V̥h, V̥ç, V̥ɸ]. Note that the present aerodynamic approach can distinguish coronal fricatives and allophones of /h/ in terms of their effects on vowel devoicing, whereas the spread-glottis approach treats them in the same way as having the feature [s.g.] (Tsuchida, 1997). Incidentally, in the present data, there are only three examples containing the sequence "plosive – V $_{[+high]}$ – /h/" which happen to allow devoicing of the accented vowel, and thus violate the constraint *V̥C $_{[+cont, -strid]}$. The other examples that have the same consonantal environment are almost exclusively unaccented, and allow devoicing of the unaccented high vowel, which satisfies this constraint.

Since there are two variants for each word, we are once again dealing with a case of free variation. A free ranking between HVD and No-Flop-Prom, as shown in (27) and (28), results in correct outputs as in (29).

(27) *V̥C $_{[+cont, -strid]}$, Max-Prom, Dep-Prom, No-Flop-Prom, *V̥ >> HVD

(28) *V̥C $_{[+cont, -strid]}$, HVD, Max-Prom, Dep-Prom, *V̥ >> No-Flop-Prom

(29) /síhai/ 'domination'

Input: /síhai/	*V̥C $_{[+cont, -strid]}$	Max-Prom	Dep-Prom	No-Flop-Prom	*V̥	HVD
a. ☞ ʃíhai						*
b. ʃi̥hái				*!		
c. ʃí̥hai	*!				*	
	*V̥C $_{[+cont, -strid]}$	HVD	Max-Prom	Dep-Prom	*V̥	No-Flop-Prom
a'. ʃíhai		*!				
b'. ☞ ʃi̥hái						*
c'. ʃí̥hai	*!				*	

As was the case in (20) and (22), different winners are produced in the upper and lower parts of the tableau by the different constraint rankings illustrated therein. Both (29c) and (29c'), which contain a voiceless accented vowel before [h], are ruled out because they violate the constraint *V̥C $_{[+cont, -strid]}$. In the upper part, (29a) is selected since it satisfies all the highest-ranked constraints, whereas (29a') is ruled out in the lower part due to the violation of HVD. In the same way, in order to predict correct outputs for words whose first variant has a voiced accented vowel and second variant has a voiceless unaccented vowel with deaccentuation (e.g., [25b] /síheN/), we need to propose another constraint ranking that pairs up with (27), producing a free ranking between HVD and Max-Prom:

(30) *V̥C $_{[+cont, -strid]}$, HVD, Dep-Prom, No-Flop-Prom, *V̥ >> Max-Prom

Note that the constraint rankings in (28) and (30) are identical to those in (19) and (21) respectively with the addition of the constraint *V̥C $_{[+cont, -strid]}$. Adding the constraint *V̥C $_{[+cont, -strid]}$ to (20) does not change the outcome for a word that does not contain an /h/ following the voiceless accented vowel, since the context specified by the constraint does not occur. (31) shows that the analysis holds with the addition of *V̥C $_{[+cont, -strid]}$ to (20):

(31) /kíka/ 'vaporization'

Input: /kíka/	*V̥C [+cont, -strid]	HVD	MAX-PROM	DEP-PROM	NO-FLOP-PROM	*V̥
a. ☞ kí̥ka						
b. ki̥ká					*!	
c. kíka		*!				
	*V̥C [+cont, -strid]	HVD	MAX-PROM	DEP-PROM	*V̥	NO-FLOP-PROM
a'. kí̥ka					*!	
b'. ☞ ki̥ká						
c'. kíka		*!				

So far, four constraint rankings have been proposed to account for the free variation observed in the interaction between accent and vowel devoicing. Table 3 in (32) shows each of the four constraint rankings and examples that can be accounted for by each ranking.

(32)

Constraint Ranking	Examples	
	Ç V[+high] Ç†	Ç V[+high] /h/
a. *V̥C [+cont, -strid], HVD, PROS-FAITH >> *V̥	kí̥ka/ʃi̥ʃa	–
b. *V̥C [+cont, -strid], HVD, MAX-PROM, DEP-PROM, *V̥ >> NO-FLOP-PROM	ki̥ká	ʃi̥hái
c. *V̥C [+cont, -strid], HVD, DEP-PROM, NO-FLOP-PROM, *V̥ >> MAX-PROM	ʃi̥ʃa	ʃi̥heɴ
d. *V̥C [+cont, -strid], PROS-FAITH, *V̥ >> HVD	* kí̥ka/*ʃi̥ʃa	ʃihai/ʃiheɴ

Table 3: Constraint rankings that have been proposed to account for free variation and the outcomes predicted by each ranking. In (32a) and (32d), PROS-FAITH stands for the three Prosodic Faithfulness constraints, which stay together in these rankings. †The following consonant can be any voiceless consonant except for /h/.

3.2.3 Vowel Length and Devoicing

So far, we have analyzed cases involving short vowels. As described above, Japanese long vowels never devoice regardless of quality, while short high vowels can devoice in certain contexts. This fact suggests that voiceless long vowels are more marked than their short counterparts. Greenberg (1969) observed that long vowels are universally less likely to devoice compared to short vowels. This tendency may be attributable to aerodynamic conditions; with long vowels, there is sufficient time to build up the necessary subglottal pressure for voicing. However, this tendency may also be related to the fact that long vowels tend to contain a tone change within the syllable (i.e., high to low or low to high according to where the long vowel is placed in the word) or a high tone throughout the syllable. As already mentioned, high-pitched vowels are unlikely to devoice; vowels manifesting a pitch change are even less likely to do so. Thus, if we adopt an aerodynamic approach to this issue, we can account for the question of long vowel devoicing in terms of vowel length and/or pitch accent, although we do not have enough evidence to decide which of the two is the more important factor.

4. CONCLUSIONS

In this paper, vowel devoicing in Tokyo Japanese has been analyzed formally using OT. Instead of the feature [s.g.] proposed by Tsuchida (1997), which is not phonologically contrastive in the Japanese grammar, and is not phonetically motivated, in the present analysis the feature [voice], which is contrastive in obstruents in Japanese, was used in such constraints as HVD, *V̥, and Ident-IO (voice). The constraints introduced in this paper such as HVD and *V̥ were aerodynamically motivated. Moreover, the rankings containing those constraints successfully predicted correct outputs in word-final position (3.2.1) and initial-accented words that show free variation concerning accent shift (3.2.2), as well as in the canonical vowel devoicing context (3.1) in Tokyo Japanese. The possible reasons that long vowels do not devoice in Japanese were also discussed in light of aerodynamic conditions. In a future study, it would be useful to test the present rankings for vowel devoicing in other Japanese dialects, including Osaka Japanese.

REFERENCES

Alderete, J. D. (1999). *Morphologically Governed Accent in Optimality Theory*. PhD dissertation. University of Massachusetts Amherst. [ROA-309, http://roa.rutgers.edu/view.php3?roa=309]

Anttila, A. (1995). Deriving variation from grammar: A study of Finnish genitives. Ms., Stanford University.

Beckman, M. E. (1994). When is a syllable not a syllable? *Ohio State University Working Papers in Linguistics* 44, 50–69.

Greenberg, J. H. (1969). Some methods of dynamic comparison in linguistics. In J. Puhvel (Ed.) *Substance and Structure of Language*. Berkeley and Los Angeles: University of California Press. 147–203.

Hirayama, T. (1960). *Zenkoku Akusento Jiten* [National Accent Dictionary]. Tokyo: Tokyodo.

Hirose, H. (1971). The activity of the adductor laryngeal muscles in respect to vowel devoicing in Japanese. *Phonetica, 23*, 156 –170.

Jaeger, J. J. (1978). Speech aerodynamics and phonological universals. In J. J. Jaeger, A. C. Woodbury, F. Ackerman, C. Chlarello, O. D. Gensler, J. Kingston, E. E. Sweetser, H. Thomson & K. W. Whistler (Eds.), *Proceedings of the Fourth Annual Meeting of the Berkeley Linguistics Society*, February 18 – 20, 1978. Berkeley, CA: Berkeley Linguistics Society, University of California. 311–329.

Kager, R. (1999). *Optimality Theory*. UK: Cambridge University Press.

McCawley, J. D. (1968). *The Phonological Component of a Grammar of Japanese*. The Hague: Mouton.

Maekawa, K. (1989). Boin no museika [Devoicing of vowels]. In M. Sugito (Ed.), *Nihongo to Nihongo Kyoiku 2: Nihongo no Onsei/Onin I* [Japanese Language and Japanese Language Education 2: Japanese Phonetics/Phonology I]. Tokyo: Meiji Shoin. 135–153.

Nihon Hoso Kyokai (ed.) (1966). *Nihongo Hatsuon Akusento Jiten* [Japanese Pronunciation Accent Dictionary]. Tokyo: Nihon Hoso.

Nitta, T. (1985). Kaga chiho ni okeru 2 mora meishi akusento no hensen [Accent change in two-mora nouns in the dialect of the Kaga region]. *Kokugogaku 140*, 103–119.

Ohso, M. (1973). A phonological study of some English loanwords in Japanese. *Ohio State University Working Papers* *in Linguistics 14*, 1–26.

Prince, A. & Smolensky, P. (1993). Optimality Theory: Constraint Interaction in Generative Grammar. Ms., Rutgers University and University of Colorado, Boulder.

Sakuma, K. (1929). *Nihon Onseigaku* [Japanese Phonetics]. References are to the 1963 reprint. Tokyo: Kazama Shoboo.

Shadle, C. H. (1997). The aerodynamics of speech. In W. H. Hardcastle & J. Laver (eds.) *The Handbook of Phonetic Sciences*. Oxford, UK: Blackwell Publishers. 33–64.

Sohn, H.-M. (1975). *Woleaian Reference Grammar*. Honolulu: The University Press of Hawaii.

Sugito, M. (ed.) (1997). *Onsei Hakei wa Kataru* [Sound Waves Tell]. Tokyo: Izumi Shoten.

Sugito, M. (ed.) (1998). *Shibata-san to Imada-san* [Mr. Shibata and Mr. Imada]. Tokyo: Izumi Shoten.

Teshigawara, M. (2001). Vowel devoicing in Tokyo and Osaka Japanese. Ms., University of Victoria.

Titze, I. R. (1992). Phonation threshold pressure: A missing link in glottal aerodynamics. *Journal of the Acoustical Society of America, 91*, 5, 2926–2935.

Tsuchida, A. (1997). *Phonetics and Phonology of Vowel Devoicing*. Ph.D. dissertation. Cornell University.

Uwano, Z. (1989). Nihongo no akusento [Accent in the Japanese language]. In M. Sugito (Ed.), *Nihongo to Nihongo Kyoiku 2: Nihongo no Onsei/Onin I* [Japanese Language and Japanese Language Education 2: Japanese Phonetics/Phonology I]. Tokyo: Meiji Shoin. 178–205.

Vance, T. J. (1987). *An Introduction to Japanese Phonology*. Albany: State University of New York Press.

Varden, J. K. (1998). *On High Vowel Devoicing in Standard Modern Japanese: Implications for Current Phonological Theory*. Ph.D. dissertation. University of Washington.

Venditti, Jennifer J. and Jan P. H. van Santen. (1998). Modeling segmental durations for Japanese text-to-speech synthesis. In *Proceedings of the 3rd ESCA Workshop on Speech Synthesis*, pp. 31-36. Jenolan Caves, Australia. [http://www.slt.atr.co.jp/cocosda/jenolan/Proc/r36/r36.pdf]

Yoshida, N. (1998). Nihongo no boin museika no onseigakuteki kenkyu [Phonetic study of Japanese vowel devoicing]. Ms., Kyoto University.

Yoshida, N. (1999). Boin museika no chiikisa [Regional differences in vowel devoicing]. Paper presented at Phonology Association in Kansai.

Yoshida, N. & Sagisaka, Y. (1990). Boin museika no yoin kenkyu [Study on factors relating to vowel devoicing]. ATR Technical Report, TR-I-0159.