Prosodic Constraint on V-to-V Coarticulation in Japanese

*Yuko Kondo** *Yumiko Arai***

Musashino Art University*,
Y & Y Phonetic Laboratory**
email: jykondo@tk.xaxon.ne.jp

ABSTRACT

Coarticulation is considered to be constrained by language specific grammar. In particular, prosodic units, which have an important function in organizing speech, are likely to affect the extent of coarticulation. The present study addresses the question of the relationship between V-to-V coarticulation and bimoraic foot in Japanese. It was shown that the foot boundary constrains the extent of coarticulation. It was also shown that both within- and across-foot, anticipatory effects are stronger than carryover effects in Japanese.

1. Introduction

The present study addresses the question of the relationship between vowel-to-vowel coarticulation and prosodic units in Japanese. Recent studies in coarticulation have shown that while coarticulation is language-universal in that it is physically constrained, its nature is also largely determined by the grammar of the language and its manifestation varies language-specifically [2, 4]. One of the most important aspects of the grammar of a language is the prosody of the language. Considering that the prosody is closely related to the organization of speech, it may be rightly presumed that it plays an important role in determining the nature and extent of coarticulation.

The smallest prosodic unit that might affect V-to-V coarticulation in Japanese is a bimoraic foot. Bimoraic foot plays an important role in the phonology of Japanese in creating hypocristics (e.g., $Akira \rightarrow Aki - chan, Megumi \rightarrow Megu - chan$) [10, 11] and in determining the accentuation of compound nouns [7, 8, 11]. Tateishi (1989) [15] also describes a bimoraic template in Japanese Musicians' language game. Also, in loan word truncation, such as /suto/ for strike and /rimokoN/ for remote control, a bimoraic foot seems to be a favourite template. Bimoraic foot also seems to play an important role in the rhythmic organization of Japanese. Sato (1994) [12] reports that Japanese 4-mora (2-foot) words exhibit the alternation of short and long syllables (See also [6]).

The present study addresses the following questions: does a bimoraic foot constrain the extent of coarticulation? More specifically, will there be any difference in the degree of V-to-V coarticulation within and across a foot boundary? Another question to be asked in the present study concerns the direction of coarticulatory effect. It has been reported that Japanese manifests stronger anticipatory than carryover effect in V-to-V coarticulation [5, 9]. The present study attempts to test the previous reports in relation to the prosodic constraint manifested by bimoraic foot.

2. Method

2.1. Subject

Six subjects participated in the present experiment. Five of them were female and the other was a male subject. Their age ranged from the mid thirties to mid forties. All of them are Tokyo dialect speakers. Five of them were born and grown up in Tokyo. One of them was originally from Nagano, a prefecture to the north of Tokyo.

2.2. Material

In order to observe the effect of the transconsonantal vowel on the two vowels /i/ and /a/ in a VbV sequence within and across bimoraic foot for both L to R (carryover) and R to L (anticipatory) effect, the following nonsense proper names were used as shown in Table 1. They are eight different names all of which sound quite alien to Japanese. The names were embedded in a carrier sentence: Kochira ga — san desu (This person is Mr. —). All the names were unaccented (LHHH). The number of moras in the sentence is even.[1] As it is reported that the coarticulatory effect may extend over a number of segments [13], the VbV sequence to be studied is embedded in the same sound context agagVbVgasa, i.e., the four segments to the right and left of the VbV are controlled. The eight experimental sentences were repeated ten times by each subject. The sentences were randomized using a random table and typed onto sheets of paper. Each subject was asked to read the sentences through 1 to 80 at a comfortable speed. Some practice session was allowed before the recording.

2.3. Analysis

The acoustic analysis was done using the signal analysis software Signalyze Version 3.0 on Macintosh Centris 660AV. The recording

[1] Tateishi proposes that foot is organized from right to left whereas Poser claims left to right formation. In order to avoid the question of the direction of foot formation, we have used four-mora (two-foot) nonsense words embedded in a carrier sentence which is made up of an even number of moras to the right and left of the test word.

Foot boundary	Direction	i	a
within	L to R	[gibi][gasa]	[giba][gasa]
		[gabi][gasa]	[gaba][gasa]
	R to L	[gibi][gasa]	[gabi][gasa]
		[giba][gasa]	[gaba][gasa]
across	L to R	[gagi][biga]	[gagi][baga]
		[gaga][biga]	[gaga][baga]
	R to L	[gagi][biga]	[gaga][biga]
		[gagi][baga]	[gaga][baga]

Table 1: The experimental material: eight nonsense proper names.

	A	B	C	D	E	F
i	26.2	14.0	39.7	-1.4	59.3	30.9
a	97.2	74.5	100.5	63.8	114.2	85.6

Table 2: The difference in F2 value (Hz) as a function of the contextual vowels (/i/ vs. /a/) for the two vowels /i/ and /a/ for each subject.

was done directly onto the computer using Apple Plain Talk microphone at the sampling rate of 22050 Hz. After the data was downsampled to 10000 Hz, the LPC analysis with the 15ms window was run to obtain the first and the second formant values. In the present study, the upper formants were not studied. The formant values used in the present study were taken at the middle of each segment.

The statistical analyses were done using the statistical package Superanova. The three-within, one-between ANOVA design was used to test the three-within variables, foot (within vs. across), direction (L to R vs. R to L) and the contextual vowel (/i/ vs. /a/) for each subject. The statistics were run separately for the vowels /i/ and /a/. Post hoc scheffe tests were also done to look at the interaction.

3. Results

3.1. Main Effect

For the vowel /a/, the main effect of the contextual vowels (/i/ vs. /a/) was significant for both F1 ($F(1,236) = 13.495$, $p = 0.0006$) and F2 ($F(1,228) = 226.802$, $p = 0.0001$). However, for the vowel /i/, the main effect of the context was significant only for F2 ($F(1,232) = 22.458$, $p = 0.0001$) and not significant for F1 ($F(1,176) = 2.391$, $p = 0.1303$). Table 2 illustrates the effect of the contextual vowels on the F2 values where significant effects were observed for both vowels /i/ and /a/. As you can see, the contextual effect is stronger on the vowel /a/ than on the vowel /i/.

3.2. Interaction

What is of interest in the present experiment is whether the effect of contextual vowels varies across the different conditions studied here, i.e.,

1. Does the effect vary within and across the foot boundary?
2. Does the effect vary depending on the direction of the effect, i.e., L to R and R to L?
3. Is there any interaction between the two conditions?

Formant	Foot	Direction	i	a
F1	within	L to R	-	32Hz (p = 0.5519)
		R to L	-	29Hz (p = 0.8009)
	across	L to R	-	24Hz (p = 0.9349)
		R to L	-	8Hz (p = 1.0000)
F2	within	L to R	55Hz (p = 0.1024)	86Hz (p = 0.0003)
		R to L	15Hz (p = 0.9881)	127Hz (p = 0.0001)
	across	L to R	41Hz (p = 0.3738)	70Hz (p = 0.0017)
		R to L	0.3Hz (p = 1.0000)	75Hz (p = 0.0001)

Table 3: The difference in F1 and F2 values of the vowels /i/ and /a/ as a function of the contextual vowels /i/ and /a/ for the following conditions: (1) L to R within foot, (2) R to L within foot, (3) L to R across foot and (4) R to L across foot.

In order to answer the questions, post hoc scheffe tests were run for the F1 and F2 of the vowel /a/ and for the F2 of the vowel /i/ where a significant main effect was observed. Table 3 shows the results. For both L to R and R to L, contextual effects were consistently stronger within a foot than across the foot boundary. However, as the p-values show, significant effects were observed only for the F2 of the vowel /a/. Therefore, the following discussion will be limited on the F2 values of the vowel /a/. The contextual effects are stronger within a foot than across the foot boundary. This trend is particularly strong for the R to L effects. For both within-foot and across-foot conditions, anticipatory effects are stronger than carryover effects. This result is consistent with the observations reported by Magen (1984) and Kondo (1995).

Table 4 shows the difference in the F2 values of the vowel /a/ as a function of contexts for each subject for the four conditions. In general the subjects show a similar trend: (1) the effects are stronger within- than across-foot, and (2) the R to L effects are stronger than L to R effects for both within- and across-foot. The exceptions to the general trend are Subject B who shows stronger L to R effect (by 6 Hz) across-foot than within-foot and Subject C who shows slightly stronger L to R effect than R to L effect across the foot boundary (by only 4 Hz). By far the notable exception is Subject E who seems to have a completely reverse trend from the other subjects. However, a closer look at the data seems to suggest that Subject E's behaviour is not totally exceptional.

In the V_1bV_2 sequence, there is a tendency for the second vowel to be lower in F2 value as shown in Table 5. This tendency is particularly strong for Subject E (the difference of about 170Hz). Even when the contextual vowel is /a/, there is difference of about 50 Hz in F2 value as a function of position across subjects (and the difference of over 200 Hz for Subject E). V_2 has consistently lower F2 values than V_1. When the contextual vowel is /i/, there are two competing effects for V_2, that is, the lowering effect due to the position and the effect of the preceding /i/ to raise the F2 value. Depending

Foot	Direction	Subject					
		A	B	C	D	E	F
within	L to R	96	47	116	82	108	71
	R to L	136	128	192	124	79	112
across	L to R	50	53	53	-4	196	65
	R to L	107	70	49	55	74	95

Table 4: The difference in F2 value (Hz) of the vowel /a/ as a function of the contextual vowels /i/ and /a/ for individual subjects for the four conditions.

	A	B	C	D	E	F
V_1	1731	1631	1802	1696	1558	1318
V_2	1677	1633	1754	1620	1388	1275

Table 5: The difference in F2 value (Hz) as a function of position in V_1bV_2 sequence.

on how these competing factors interact with one another, we get different patterns as illustrated in the case of Subject E (See Figure 1). For Subject E, when we look at the F2 values in the context of /i/, the following /i/ seems to exert stronger effect than the preceding /i/. Her F2 value is highest when the vowel is followed by a foot internal /i/ just like the rest of the subjects. However, a large drop in F2 observed for V_2 in the /a/ context masks the effect of the contextual /i/ when only the difference in F2 value is considered.

We may summarize the trend observed across subjects as follows:

1. The F2 value decreases in the following order: followed by foot-internal /i/ (1694Hz) $>$ followed by foot-external /i/ (1642Hz) $>$ preceded by foot-internal /i/ (1600Hz) $>$ preceded by foot-external /i/ (1580Hz).
2. In the context of /a/, the second vowel in the VbV shows a similar drop in F2 value for both within- and across-foot (when Subject E is excluded from the data, the same sort of drop is still observed though the extent of drop is much lower).

The two factors, the effect of /i/ to raise F2 values and the lowering of F2 due to position (presumably, this factor works in the context of /i/ as well as in the context of /a/), seem to interact with one another to produce the observed effect in the direction of coarticulatory effect: R to L effect seems to be stronger than L to R effect for most speakers. However, we must take into account the drop in F2 value in the V_2 position though it may be small for most of the speakers. That the contextual /i/ exerted stronger effect when it is foot-internal than when it is foot-external suggests that bimoraic foot constrains the extent of coarticulation.

4. Discussion

In the present study, the Japanese vowel /i/ showed stronger resistance to the coarticulatory effects and showed less variability compared to the vowel /a/. This sort of asymmetry in the degree of variability between /i/ and /a/ is not surprising as Bates (1995) reports that vowels exhibit their own inherent variability as a function of context [1]. However, /i/ ranks higher in the context-sensitive variability compared to /a/ in her British English data. Magen (1984) reports that while asymmetry in context-sensitive variability was observed between /i/ and /a/ in Japanese, such asymmetry was not observed for English. On the other hand, Farnetani (1990) [3] reports that the vowel /i/ in Italian is less susceptible to coarticulatory effects than /a/ as in the case of Japanese. It seems that this asymmetry in coarticulatory sensitivity between /i/ and /a/ is language-specific.

The present experiment showed that V-to-V coarticulation observed in the F2 value of the Japanese vowel /a/ is stronger within foot than across foot. That is, bimoraic foot does seem to constrain the degree of coarticulation. It also showed that anticipatory effect was stronger than carryover effect. On the other hand, it was observed that the two effects, the F2 raising effect of the contextual /i/ and the F2 lowering effect of the second position in VbV sequence, interacted. The precise reason for this lowering in F2 is not clear. The consonant /b/ is known to lower the F2 values [14]. As there is a syllable or mora boundary between the first V and bV (V /bV), it is plausible that the medial /b/ exerts stronger effect on the second V than on the first V. What is of importance here is that coarticulation seems to be a result of the complex interaction of multi-levelled constraints, and a series of carefully planned experiments are necessary to uncover the precise function and the extent of influence of each constraint. Different surface patterns may result from different interactions of the same constraints. Some constraints may work more strongly for some speakers or languages. This concept of constraint interaction may explain speaker variability and language variability in coarticulation.

REFERENCES

[1] S.A.R. Bates. Acoustic characterization of schwa: A comparative study. In *Proceedings of ICPhS 95*, 1995.

[2] H. Clumeck. Patterns of soft palate movements in six languages. *Journal of Phonetics*, 4:337–351, 1976.

[3] E. Farnetani. V-C-V lingual coarticulation and its spatiotemporal domain. In W.J. Hardcastle and A. Marchal, editors, *Speech Production and Speech Modeling*, pages 93–130. Kluwer Academic Press, Dordrecht, 1990.

[4] W.J. Hardcastle. Constraints on coarticulatory processes. In D. Crystal, editor, *Linguistic Controversies: Essays in Linguistic Theory and Practice in Honour of F.R. Palmer*, pages 33–49. Edward Arnold, London, 1982.

[5] Y. Kondo. *Production of schwa by Japanese speakers of English: a cross-linguistic study of coarticulatory strategies*. PhD thesis, University of Edinburgh, 1995.

[6] Y. Kondo and Y. Arai. Prosodic constraint on V-to-V coarticulation in Japanese: a case of bimoraic foot. In The Phonological Society of Japan, editor, *Phonological Studies*, pages 75–82. Kaitaku-sha, Tokyo, Japan, 1998.

[7] H. Kubozono. Constraint interaction in Japanese phonology: Evidence from compound accent. *Phonology at Santa Cruz*, 4:21–38, 1994.

[8] H Kubozono. Lexical markedness and variation: A non-derivational account of Japanese compound accent. *WCCFL*, 15:273–287, 1997.

[9] H. Magen. Vowel-to-vowel coarticulation in English and Japanese. *Journal of the Acoustical Society of America*, 75, Suppl.:1:S11, 1984.

[10] A.R. Mester. Patterns of truncation. *Linguistic Inquiry*, 21-3:478–485, 1990.

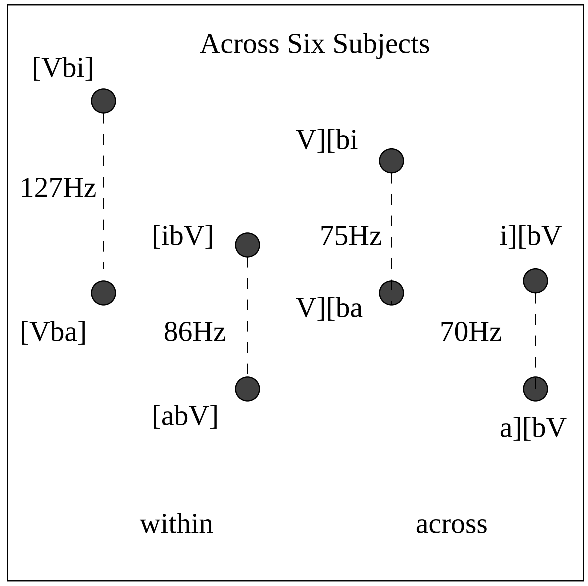
[11] W.J. Poser. Evidence for foot structure in Japanese. *Language*, 66:78–105, 1990.

[12] H. Sato. Bimoraic rhythm and accentuation in Japanese. In *The twelfth national conference of the English Linguistic Society of Japan*, Tokyo, 1994.

[13] D.J. Sharf and R.N. Ohde. Physiologic, acoustic, and perceptual aspects of coarticulation: implications for the remediation of articulatory disorders. In *Speech and Language: Advances in Basic Research and Practice*, volume 5, pages 153–247. Academic Press, 1981.

[14] K.N. Stevens and A.S. House. Perturbation of vowel articulations by consonantal context: an acoustical study. *Journal of Speech and Hearing Research*, 6:111–128, 1963.

[15] K. Tateishi. Theoretical implications of the Japanese musicians' language. *WCCFL*, 8:384–398, 1989.



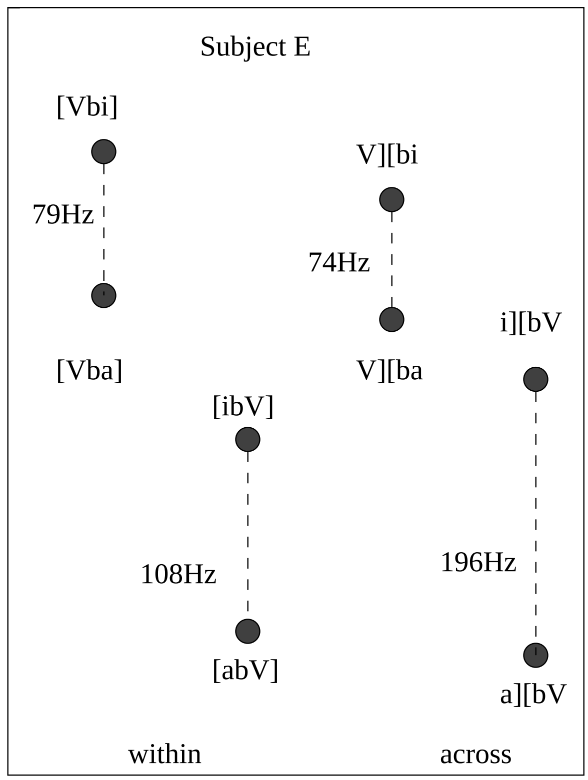


Figure 1: Schematic figures showing the distribution of the F2 values of the vowel /a/ in the $foot \times direction(position)$ conditions across subjects and for Subject E. (See also Tables 3 and 4.)