

Compensatory Lengthening by British Learners of Japanese

NAGAI Katsumi*

Key words: **mora, duration, compensatory lengthening**

Every phoneme has its own intrinsic segmental duration, which varies in the environment. This experiment examines the effect of voicing second consonants in C1V1. C2V2 test words spoken by Japanese native speakers and British learners of Japanese at two levels. Compensatory effect, which is seen between the voiced consonant and the preceding vowel, occurs beyond CV boundaries. Elementary learners can be distinguished because they have little effect of durational compensation. The result implies a possibility of applying the effect to evaluating the learners' achievement of second language learning.

INTRODUCTION AND PRECEDING STUDIES

Duration, as a phonetic and physical unit of time of speech event, needs to be distinguished from the phonological length. In this paper, duration is defined as a speaker's time between start-point and end-point of utterance, which is considered to be free from the divergence among listeners. Because everybody has their own default tongue setting, it takes more time for a tongue move to the destination when the tongue deviates from the natural position for a schwa. Of course, each segmental duration becomes longer when the speaker speaks at a slow tempo. A stressed syllable also has longer duration than the other unstressed syllables in an English word (Fry 1955).

It is said that every phoneme of languages has its own segmental duration

* 長井克己: Department of Applied Linguistics, Edinburgh University, Scotland.

This research at Edinburgh University was supported by Rotary Foundation Grant 1996. I am also grateful to all staff at the Department of Linguistics and Applied Linguistics, especially to Dr. Turk, Dr. Sorace, and Mr. K. Mitchell for their kind advice. All errors and misunderstandings are mine. Please write to nagai@holyrood.ed.ac.uk or nagai@lisa.lang.osaka-u.ac.jp

(Lehiste 1970: 18ff.). First, English tense vowels and diphthongs are usually longer than their corresponding lax vowels as observed between /i/ and /I/. Second, differences in duration are reported as a function of place of articulation. Klatt (1976) clarifies that bilabial stops are typically slightly longer in duration than alveolars and velars. Third, voiceless obstruents are longer than voiced sound. Klatt (1976) says voiceless fricatives are about 40 ms longer in duration than the corresponding voiced fricatives. He also shows that the voiceless fricative /s/ in “soo” is inherently longer than the voiced fricative /z/ in “zoo.” Borden *et al.* (1994: 142) explains: “while these duational differences are found across language, suggesting that it is basically conditioned by physiology, English shows very large differences in vowel duration before voiced and voiceless consonants, suggesting a learned overlay.”

In the case of Japanese, Sugito (1996: 274) clarifies that a native speaker’s second unvoiced consonant of “*tata*” is longer than the second voiced consonant of “*dada*,” while the reverse is true for Chinese speakers. Research on Spanish has also shown a small average difference (18 ms) between vowels preceding voiced and voiceless consonants. Campbell’s report (1992: 213) states that Japanese vowels are, contrary to English, shorter when they are followed by voiced consonants. This inconsistency implies that the effect of voicing a consonant on the preceding vowel is not an innate factor. Therefore, it is necessary to test compensatory lengthening by both native and nonnative speakers. Would it be reasonable to assume an inherent relative difference of segmental duration between the utterances of native Japanese speakers and British learners of Japanese?

The Experiment

1 Aim

The aim of this experiment is to clarify whether voicing of the second consonants of Japanese CV.CV words spoken by both English and Japanese speakers have an effect on each segmental duration. It is hypothesized that the intrinsic effect of voicing in Japanese speech sound is consistently affected by the difference of learners’ achievement level. Is the duration of voiceless consonants by Japanese speakers consistently longer than that of voiced consonants? Does the vowel in the first CV sequence become longer across the moraic boundary when the second consonant is [+voiced]?

2 Design

Pairs of test words are read by subjects to investigate whether the effect of voicing the second consonant is at work. Supposing both native Japanese speakers and British learners of Japanese have their own intrinsic segmental duration of speech sound, it is plausible that the difference of their first language has an effect on the segmental duration and compensation of the change. If English speakers’ voicing of the second consonant gives a different effect from that of

Japanese speakers', it would be a sign of their phenetically intrinsic segmental duration being overwhelmed by the effect of voicing.

The null hypotheses in this experiment are (1) H_0 : Native speakers' and advanced learners' duration of a voiceless segment in a Japanese sentence is as long as that of their voiced counterparts, and (2) H_0 : Native speakers' and advanced learners' vowel duration preceding the voiced consonant does not change because the compensatory effect occurs only within a single moraic CV boundary. Data from the advanced learners are expected to fall between native speakers and elementary learners.

3 Materials

The following sets of test words were read by subjects. The words shown below without glosses are nonsense words. Although Sugito (1989: 172) reports that Japanese syllables do not change in duration even if the test words in a frame sentence have different pitch patterns, all the test words were read with a HL pitch pattern.

<i>kaka</i>	<i>kata</i> (shoulder)	<i>kasa</i> (umbrella)
<i>kaga</i> (place name)	<i>kada</i> (place name)	<i>kaza</i>

The six words above were embedded in the frame sentence:

<i>Kore-ga</i>	. . . <i>desu</i> .
"This-NOM is

It must be mentioned that some of the elementary learners misread the test sentences during the first rehearsal time. All the British subjects have learned Japanese *kana* alphabets, however, the elementary learners sometimes seem to have trouble reading the list because Japanese voiced sound is written in a character very similar to its voiceless counterpart. That is, just adding double dots (¨) makes a difference between voiced (か *ka*) and voiceless (か¨ *ga*) segments in Japanese. If all the test words are written in Roman transcriptions, on the contrary, native Japanese speakers would feel unnatural and have difficulty when reading the sentences. This fact implies a need for research without using alphabets or *kana*, such as close examination of spontaneous speech outside the recording room.

A set of test words which are comprised of only nonsense words might be more suitable, however, Klatt (1976) concludes that the similarities between nonsense syllable studies and spontaneous speech are greater than the differences. The reading lists were written in *hiragana*. Total number of utterances amounted to 6 words \times 5 times \times 12 subjects = 360.

4 Subjects

Subjects in this experiment are four native speakers of Japanese, four British elementary learners of Japanese, and four British advanced learners of Japanese. All the native Japanese are speakers of standard Tokyo dialect, and are all language teachers in Japan.

The British elementary learners of Japanese are sophomores at the University

of Edinburgh. They have studied Japanese one and half years, including four lectures on basic Japanese grammar and two follow-up tutorials per week. They have no experience in studying abroad.

The British advanced learners are senior students at the same university and lived in Japan for one year as exchange students studying Japanese language. The curriculum they finished includes Japanese grammar, translation into and from Japanese, conversation and discussion, and kanji. Although their vocabulary in Japanese is somewhat limited and some unnatural accentuation still remains, they have little difficulty in making themselves understood in Japanese.

5 Procedure

Each sentence was read five times after a few minutes of practice. Their utterances were recorded using a sensitive condenser microphone (Senheisser MKH-815). The signal was sent to a microphone amplifier (Soundcraft 200B) and recorded digitally on a DAT recorder (SONY PCM-2700A) with 16 bit/44.1 KHz sampling. The recordings were analyzed on a UNIX workstation (Sun Sparc Station 5) with a D/A and A/D conversion board. The duration of the segmental units of words were measured by wide-band spectrograms and time domain waveforms on a VDT screen of X-Waves analyzer with digitization of 16 KHz sampling frequency with aid of intensity curves and time domain curves (Figures 1 and 2). The criteria used for pinpointing cursors in the display are shown in Table 1. When more than two burst

Table 1 Measurement of Duration

1.	Consonant 1 (/k/) closure duration: the interval between the offset of energy in the formants of the preceding vowel (/a/) and the release burst spike.
2.	Voice Onset Time (VOT) of Consonant 1 (/k/): from the first release of bursts to the first visible glottal pulse of the following vowel (/a/). In case that the following vowel is devoiced, the VOT becomes the entire aspiration-filled interval between the two /k/ closures.
3.	Vowel 1 (/i/ or /a/) duration: from the first glottal pulse to the closure of the following consonant.
4.	Consonant 2 (/k/, /g/, /t/, /d/, /s/, or /z/) duration: from the offset of energy in the formants of Vowel 1 (/a/) to either the release burst of Vowel 2 (for /k/ and /t/) or the onset of energy in the formant structure of Vowel 2 (for /s/, /z/, /d/, and, /k/ or /t/ without release burst evidence. See Figure 1 and Figure 2.) In cases that a trace of nasalized /g/ is detected, the duration of the nasal constriction is also measured.
5.	Voice Onset Time for Consonant 2 (/k/ and /t/): from the beginning of the first release burst on the initial stop to the first visible striation representing glottal pulsing. In case that very weak pulses called "edge vibration" (Lisker and Abramson 1967) are observed, they are ignored as inaudible signals.
6.	Vowel 2 duration: from the first glottal pulse following the release of Consonant 2 to the closure for the /d/ of the carrier word <i>desu</i> .

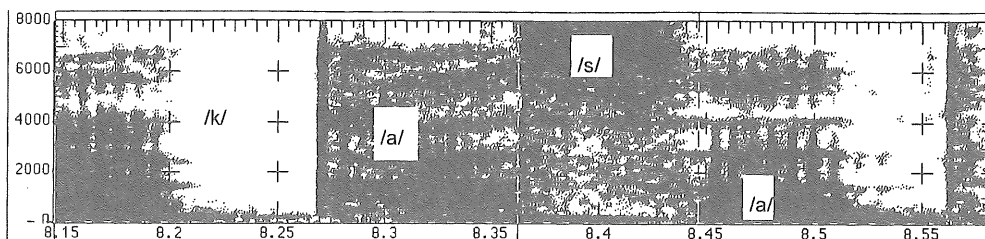


Fig. 1 Measurement of *kasa* by a native speaker of Japanese

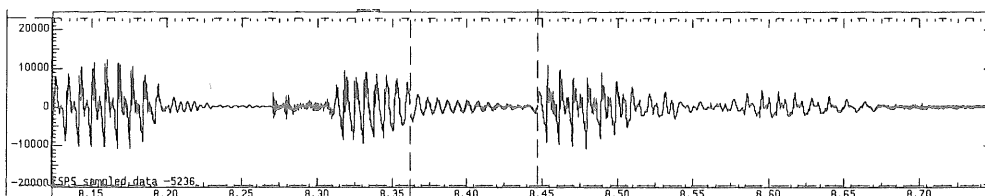


Fig. 2 Waveform of Figure 1

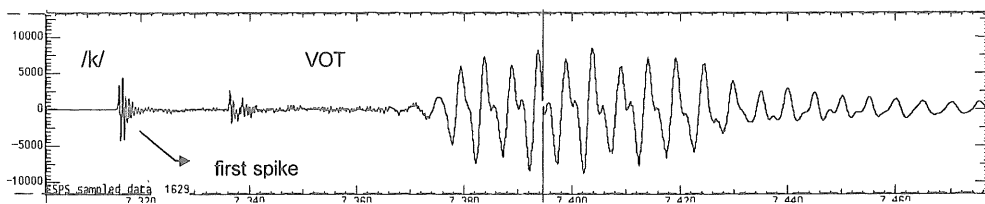


Fig. 3 An example of two burst spikes

spikes were observed before the offset of the consonantal closure, the first one was used to measure the voice onset time as shown in Figure 3. The cursors in a display window were reset every time of pinpointing (Ohala and Lyberg 1976).

The independent variable 'GROUPS' was at three levels (native speakers, elementary learners, and advanced learners), while 'VOICED' was at two level (+/-voiced) when repeated measure ANOVA was executed. The VOTs are included in V1 or V2 when data is statistically tested because they are considered to vary consistently across the place of articulation of the stop sound (Port and Rutunno 1979).

6 Results

Mean duration and the standard deviations are shown in Table 2. Each column is means pooled across the subjects. The results are plotted in Figures 4, 5, and 6. The most eminent feature is that the native speakers' overall word duration is similar despite its segmental varieties. The total word duration of

kaga is slightly longer than the rest of test words. However, almost all of the native speakers keep their duration consistent in spite of the variety of segmental duration. On the one hand, it means that if the number of mora is the same, the duration of the word looks almost the same because of its mora-timing. On the other hand, word duration by elementary learners looks much more deviated than that of native speakers. Total word duration of the elementary learners becomes longer when the test words include voiced C2. It implies that isochronal mora timing is not realized correctly when elementary learners speak Japanese words. Advanced learners fall between native speakers and elementary learners when it comes to total word duration.

Repeated measure ANOVA reveals that voiced C2 duration by native speakers is shorter than voiceless C2 ($F_{(1,22)}=11.42$ $p<0.005$), but that voiced C2 duration by elementary learners is not shorter than their voiceless counterparts ($F_{(1,22)}=0.15$ n.s.). Advanced learners fall between native speakers and elementary learners ($F_{(1,22)}=3.29$ $p=0.083$). This result enables the null hypothesis (1) to be rejected. While native speakers' voicing of C2 makes itself shorter than voiceless C2, the voicing of C2 makes the preceding V1 longer ($F_{(1,22)}=17.26$ $p<0.001$). The same is true for advanced learners ($F_{(1,22)}=11.09$ $p<0.005$), but this compensatory effect is not seen in elementary learners' utterance ($F_{(1,22)}=0.13$ n.s.). Then the null hypothesis (2) needs to be rejected too.

The subject groups have a significant main effect on all total word duration

Table 2 Mean Segmental Duration (S.D.) in ms.

	C1 (/k/)	VOT1	V1 (/a/)	C2	VOT2	V2 (/a/)	V1+VOT	V2+VOT	Word
Native speakers									
<i>kaka</i>	70(25)	34 (9)	60(12)	61(17)	25 (5)	80(10)	93 (9)	105(11)	329(42)
<i>kaga</i>	78(35)	40(16)	76(13)	41 (9)		99(17)	116(14)	99(17)	335(55)
<i>kata</i>	75(27)	32(10)	58(11)	59(15)	20 (4)	81(13)	90 (9)	101(13)	326(46)
<i>kada</i>	85(29)	40(16)	66(14)	38(18)		97(20)	106(16)	97(20)	326(49)
<i>kasa</i>	78(30)	33(10)	60(13)	81(16)		78(18)	93(14)	78(18)	330(51)
<i>kaza</i>	80(27)	36 (9)	71(10)	55(20)		92(15)	107(11)	92(15)	334(53)
Elementary learners									
<i>kaka</i>	111(44)	54(12)	90(15)	76(28)	51(38)	142(34)	145(12)	193(40)	525(68)
<i>kaga</i>	147(41)	62(17)	87(17)	134(47)		188(25)	150(25)	188(25)	619(57)
<i>kata</i>	133(56)	58(15)	92(13)	105(37)	39(19)	148(41)	150(22)	187(33)	576(105)
<i>kada</i>	161(60)	55(18)	95(15)	113(49)		174(32)	150(27)	174(32)	598(75)
<i>kasa</i>	124(38)	62(16)	109(67)	153(68)		150(34)	171(70)	150(34)	597(103)
<i>kaza</i>	130(49)	68(23)	110(20)	115(50)		186(53)	178(28)	186(53)	609(102)
Advanced learners									
<i>kaka</i>	115(47)	42 (6)	69(12)	52(14)	31 (8)	82 (9)	111(14)	113(12)	391(52)
<i>kaga</i>	109(43)	47(11)	83(17)	39(19)		120(31)	130(18)	120(31)	397(61)
<i>kata</i>	96(28)	43 (9)	60(12)	56(17)	23 (5)	75 (8)	102(14)	97(10)	352(39)
<i>kada</i>	118(35)	56(14)	74(20)	50(30)		105(22)	130(20)	105(22)	404(82)
<i>kasa</i>	105(39)	51(16)	68(12)	91(13)		85(20)	119(19)	85(20)	399(58)
<i>kaza</i>	107(40)	57(16)	76(15)	65(11)		93(22)	134(25)	93(22)	399(50)

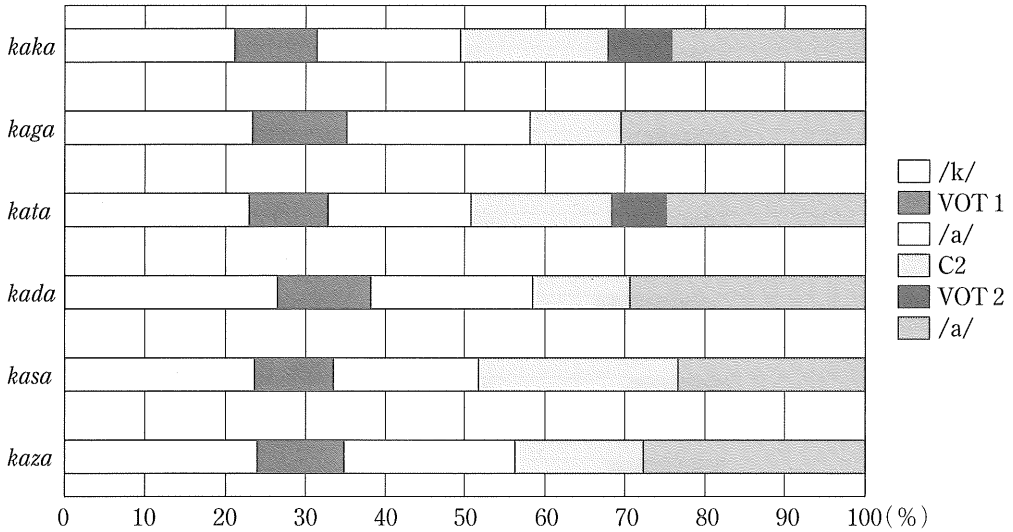


Fig. 4 Segmental Duration by Native Speakers

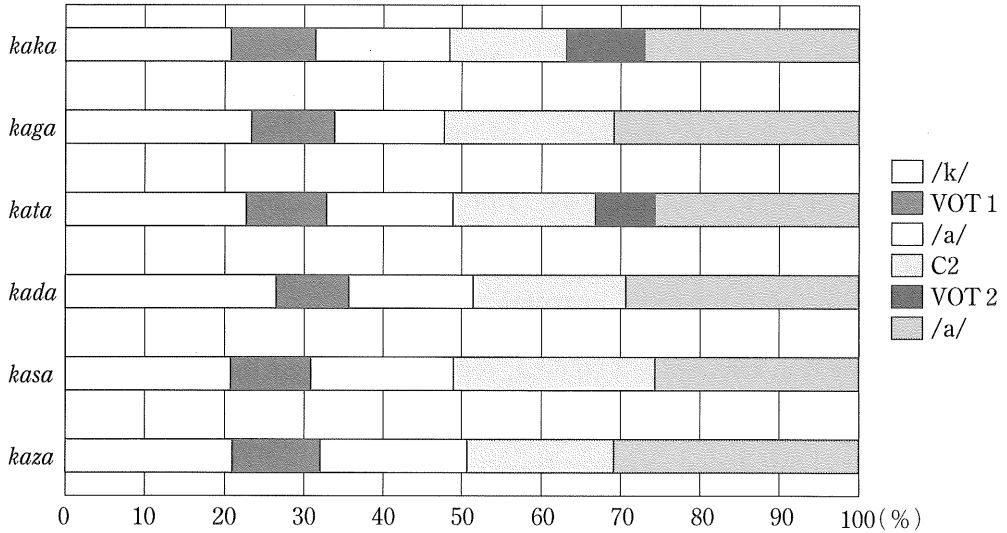


Fig. 5 Segmental Duration by Elementary Learners

($F_{(2,69)}=141.57$ $p<0.001$) across all test words probably because elementary learners take nearly twice as long to utter the test words than native speakers. The total word duration of each group is, however, not affected by the distinction of [+/-voiced] features (native speakers: $F_{(1,22)}=0.09$ n.s., advanced learners: $F_{(1,22)}=1.07$ n.s.). It implies that the total duration of Japanese words could be used, too, as a scale of learners' achievement in second language

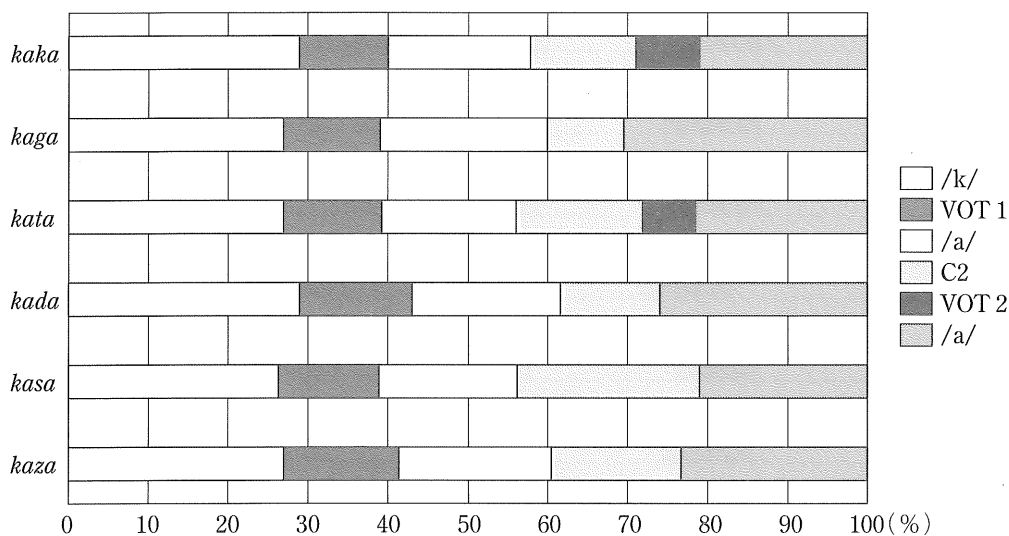


Fig. 6 Segmental Duration by Advanced Speakers

phonology.

The /k/ and /t/ in C2 position are different from each other in point of the place of articulation. This difference gives no effect on the total word duration ($F_{(1,14)}=0.04$ n.s.) of native speakers. In this distinction of /k/ and /t/, elementary learners ($F_{(1,14)}=0.19$ n.s.) and advanced learners ($F_{(1,14)}=0.39$ n.s.) showed no difference of total word duration either. Another pair of /s/ and /z/ in C2 position differ in their manner of articulation. They also exhibit no effect on the total word duration (native speakers: $F_{(1,14)}=0.01$ n.s., elementary learners: $F_{(1,14)}=0.19$ n.s., and advanced learners: $F_{(1,14)}=0.78$ n.s.). The results above reconfirm the findings by Port *et al.* (1987) with native Japanese subjects. Then it seems that effect of C2 voicing, regardless of their different articulation, would be a suitable scale to evaluate learners' SLA phonology, because effects of the manner of articulation in Japanese language are, as Campbell (1987) and Sugito (1989) pointed out, not very large.

DISCUSSION AND CONCLUSION

The intrinsic difference in duration between voiced and voiceless consonants plays a key role in the experiment. It has been reported in Port *et al.* (1980)¹

¹ They reported strong correlation calculated only from the data of which absolute duration of C1V1 matched with its C2V2 duration. It is not clear whether this kind of data picking will not skew the data's randomness for a statistical test. Native speakers' Pearson's correlation coefficient across all data was not significant ($r=-0.1369$, $p=0.136$) in this experiment.

and Campbell (1992) that voiced consonants are intrinsically short and that a vowel becomes longer when it precedes the voiced consonant. The result above revealed that the same kind of compensatory segmental adjustment is observed in the speech of native speakers and advanced learners of Japanese. While the Japanese voiceless obstruent C2 read by English speakers is longer than voiced ones, British elementary learners of Japanese could not share this effect perhaps because of their insufficient phonological system, which is expected to cause the lengthening of devoiced consonants. In other words, learners can be classified from a viewpoint of whether they can show the universal-looking lengthening effect of voicing. It might be one of the elementary learners' targets to achieve more natural utterances.

A more conspicuous fact would be that a compensatory effect occurs beyond the CV boundary of the Japanese C1V1. C2V2 word. That is, C2 segmental duration gives effect on the preceding V1 duration across the moraic CV boundaries. This compensation by C2 for the duration of preceding V1 is observed only in native Japanese speakers' and advanced learners' utterances. It implies that the compensatory effect across the boundary of two morae might be also one of the important factors of natural speech in Japanese because V1 duration by elementary learners does not vary by the effect of its following C2 duration. It may support the existence of a unit of two morae, or bimoraic foot, which explains Japanese compensatory effect more well formedly.

It is true, in this experiment, that a unit of Japanese CV morae is also unit of CV syllables at the same time. Though discussion on phonological segmentation of interlanguage is beyond the scope of this short paper, Ôtake's experiment (1992) of target monitoring tasks also hints at the coexistence of syllables and morae in Japanese phonology. Apart from the dispute of whether it is a language-universal effect or not, language teachers and researchers need to keep studying what prevents the elementary learners' compensatory lengthening and how their first language interacts with their interlanguage phonology.

Traditional pedagogy has claimed that the basic phonological unit of Japanese is moraic (C)V sequences, which have constantly equal duration. However, teachers of Japanese need to know a more important fact: that compensatory timing control of Japanese is realized not only inside of a single CV sequence but also across the adjacent CV units to keep the total word duration consistent.

BIBLIOGRAPHY

- Beckman, M. 1982. 'Segmental duration and the 'mora' in Japanese.' *Phonetica* 39: 113-35.
- Beckman, M. and Atsuko Shoji. 1984. "Spectral and Perceptual Evidence for CV Coarticulation in Devoiced /si/ and /syu/ in Japanese." *Phonetica* 41: 61-71.
- Borden, G. J., K. S. Harris and L. J. Raphael. 1994. *Speech science primer*, third edition. Baltimore: Williams & Wilkins.

- Blumstein, S. E. and K. N. Stevens. 1979. "Acoustic invariance in speech production: Evidence from measurements of the spectral characteristics of stop consonants." *Journal of the Acoustical Society of America* 66: 1001-17.
- Campbell, N. 1992. 'Speech Timing in English and Japanese.' Proceedings for International Symposium on Japanese Prosody. Nara, Japan.
- Clark, J. and C. Yallop. 1995. *An introduction to phonetics and phonology*, second edition. Oxford: Blackwell.
- Cutler, A., J. Mehler, D. Norris and J. Segui. 1986. 'The syllable's roll in the segmentation of French and English.' *Journal of Memory and Language* 25: 385-400.
- Flege, J. E. and J. Hillenbrand. 1984. 'Limits on phonetic Accuracy on Foreign Language Speech Production.' *Journal of the Acoustical Society of America* 76 (3): 706-21.
- Fowler, A. 1980. "Coarticulation and theories of extrinsic timing." *Journal of Phonetics*, 8: 113-33.
- Fry, D. B. 1955. "Duration and intensity as Physical Correlates of Linguistic Stress." *Journal of the Acoustical Society of America* 27: 765-68.
- Ganong, W. F. and R. J. Zatorre. 1980. 'Measuring phoneme boundaries in four ways.' *Journal of the Acoustical Society of America* 68: 431-39.
- Han, M.S. 1962. 'The Feature of Duration in Japanese.' *Onsei no Kenkyū* 10: 65-80.
- Hoequist, C. Jr. 1983. "Syllable Duration in Stress-, Syllable- and Mora-Timed Languages." *Phonetica* 40: 203-37.
- Kaiki, N. and Y. Sagisaka. 1992. 'The control of segmental duration is speech synthesis using statistical methods.' In *Speech perception, production and linguistic structure*, ed. Tōkura, Amsterdam: IOS Press.
- James, A. and J. Leather, ed. 1986. *Sound patterns in second language acquisition*. Dordrecht: Foris.
- Klatt, D. H. 1975. "Vowel lengthening is Syntactically Determined in a Connected Discourse." *Journal of Phonetics* 3: 129-40.
- Klatt, D. H. 1976. "Linguistic uses of segmental duration in English: Acoustic and perceptual evidence." *Journal of the Acoustical Society of America* 59 (5): 1208-21.
- Ladefoged, P. 1993. *A course in phonetics*, third Edition. New York: Harcourt Brace Jovanovich.
- Ladefoged, P. and I. Maddieson. 1996. *The sounds of the world's language*. Oxford: Blackwell.
- Laver, J. 1994. *Principles of phonetics*. Cambridge: Cambridge University Press.
- Lehiste, I. 1970. *Suprasegmentals*. Cambridge, Mass.: MIT Press.
- . 1972. "The timing of utterances and linguistic boundaries." *Journal of the Acoustical Society of America* 51: 2018-24.
- Lisker, L. and Abramson, A. 1967. "Some effects of context on voice onset time in English stops." *Language and Speech* 10: 1-28.
- Mehler, J., J. Dommergues, U. Frauenfelder, and J. Segui. 1981. 'The syllable's role in speech segmentation.' *Journal of Verbal Learning and Verbal Behavior* 9: 295-302.
- Meyer, A. S. 1992. 'Investigation of phonological encoding through speech error analysis: Achievements, limitations, and alternatives.' *Cognition* 42: 181-211.

- Morton, J. and C. Frankish. 1976. "Perceptual centres (P-centers)." *Psychological Review* 83: 405-8.
- Ohala, J. and S. Lyberg. 1976. "Comments on 'temporal interactions within a phrase and sentences.'" *Journal of the Acoustical Society of America* 59: 990-92.
- Pierrehumbert, Janet. 1994. 'Syllable structure and word structure: a study of triconsonantal clusters in English.' In P. A. Keating, ed. *Phonological structure and phonetic form: Papes in laboratory phonology III*. Cambridge: Cambridge University Press: 168-90.
- Port, R. F., J. Dalby, and M. O'Dell. 1987. 'Evidence for mora timing in Japanese.' *Journal of the Acoustical Society of America* 81 (5): 1574-85.
- Port, R. F., S. Al-Ani, and S. Maeda. 1980. 'Temporal compensation and universal phonetics.' *Phonetica* 37: 235-52.
- Port, R. F. and R. Rotunno. 1979. "Relation between voice-onset time and vowel duration." *Journal of the Acoustical Society of America* 66: 654-62.
- Takagi, Naoyuki and V. Mann. 1994. "A perceptual basis for the systematic phonological correspondences between Japanese load (*sic.*) words and their English source words." *Journal of Phionetics* 22: 343-56.
- Vance, T. J. 1987. *An introduction to Japanese phonology*. New York: State University of New York Press.

大学と科学公開シンポジウム組織委員会 (1993) 「国際化する日本語」, クバプロ.

大竹孝司 (1992) 『日本語音声の知覚上の単位: 音節とモーラ』, 原口編 「日本語のモーラと音節構造の研究 2」, 文部省重点領域研究「日本語音声」研究成果刊行書.

杉藤美代子 (1989) 『日本語と英語のアクセントとイントネーション』, 杉藤編 「日本語と日本語教育第 2 巻」, 東京: 明治書院.

——— (1996) 「日本語の音」, 大阪: 和泉書院.