Rate induced resyllabification and sonority scale in Japanese

Mafuyu Kitahara*, **Tadao Murata**† and **Haruka Fukazawa**†

* Yamaguchi University, 2-16-1 Tokiwadai, Ube, Japan
† Kyushu Institute of Technology, 1-1 Sensui-cho, Tobata-ku, Kitakyushu, Japan

ABSTRACT

The present study investigates the rate induced resyllabification process for various consonants and vowels in Japanese. The purpose is to find whether there are articulatory correlates of sonority. The switching point from a VC to a CV syllable was the main dependent variable. Results of an production experiment show that the switching point analysis closely replicates the scale given by sonority.

1 INTRODUCTION

Sonority is a frequently invoked, but not tightly defined property of (classes of) segments in the phonological literature. In particular, the order of segments on a sonority scale and its universality has been a controversial topic. Still, it is considered to be useful because the sonority sequencing has been a guiding principle for a syllabification process. Based on binary distinctive features, a working universal sonority scale is proposed [1]. Those relevant to the present study are summarized in (1)–(3).

$$\textit{voiceless stop} < \textit{voiced stop} \quad (1)$$

$$\textit{stop} < \textit{fricative} \quad (2)$$

$$\textit{obstruent} < \textit{sonorant} \quad (3)$$

The notion of sonority is loosely connected to perceptual "loudness", but its articulatory correlates have not yet been clearly determined. At a certain level of abstraction, articulatory "effort" might constitute an inverse scale of sonority. However, the degree of effort is not easy to measure in speech. Low-level kinematic parameters, such as the mass and the velocity of articulators may differ in various ways for each consonant.

The present study examines whether there is an articulatory correlates of sonority by looking at a rather high-level variable: a switching point from a VC to a CV syllable when the VC syllable is repeated at an increasing rate. This rate induced resyllabification is originally noted by R. H. Stetson and replicated in various formats [2, 3, 4, 5]. These previous studies mainly focussed on overall prosodic and dynamic properties of syllabification to reveal the interplay between production and perception. However, consonants used in their experiments are limited to only [p] and [b]. It is worth exploring other types of consonants in this rate scaling paradigm.

Thus, the segmental material in the present study consists of twelve consonants and two vowels. Our purpose is to examine whether the VC-to-CV switching occurs in these segmental combinations and, if it does, to find where in the rate scale the switching occurs.

2 METHODS

Eight male subjects (age: 18–21) participated the experiment. They were undergraduate students at Kyushu Institute of Technology (KIT) and the experiment was carried out as a part of an English class. Most of them speak Northern-Kyushu varieties of Japanese. Although they all had more than 6 years of English experience in school, their proficiency level was generally poor. This led to a serious difficulty in articulating a VC syllable for them. As is well known, English loans in Japanese have abundant epentheses. For example, 'strike' becomes [s**u**t**o**raik**u**] due to constraints on syllabic organization in Japanese. As expected, subjects often epenthesized [u] or [o] after the syllable final C in the present experiment. The limitation on the data analysis due to epenthesis will be discussed in the subsequent sections.

Segmental material of the VC syllable consists of twelve consonants [p,t,k,b,d,g,f,s,z,m,n,l] and two vowels [i,a]. Subjects were divided into two groups of four people. Group1 pronounced only [iC] syllables while group2 pronounced only [aC] syllables. They were all instructed to entrain to an electronic metronome whose tempo increased from 300msec to 100msec (onset-to-onset interval) in each trial. There are total 66 beeps in a single trial. First 6 beeps kept 300msec interval and next 50 beeps decreased their intervals by 4msec. Final 10 beeps kept 100msec interval. Each metronome beep was a 400Hz pure tone and had a 50msec duration.

Recordings were done in a language laboratory facility at KIT. Metronome beeps were presented to subjects via headphones. Speech was recorded on the left channel and the metronome beep was simultaneously recorded on the right channel of an analogue cassette tape in order to examine the performance of subjects' entrainment. There was one practice session before the experiment. Then, a randomized list of 12 VC syllables was given to each subject. Recorded sounds were digitized at 22kHz/16-bit quality and inspected visually on a spectrogram (*praat* 4.29 on Linux) for measurements.

A schematic illustration of the recorded sounds is given in Figure 1. A typical token in a trial has three measurement points: consonant release, voice onset, and the beep onset nearest to the voice onset. Three durational intervals can be calculated from those points: beep to voice (**B2V**), consonant to vowel (**C2V**), and vowel to consonant (**V2C**).

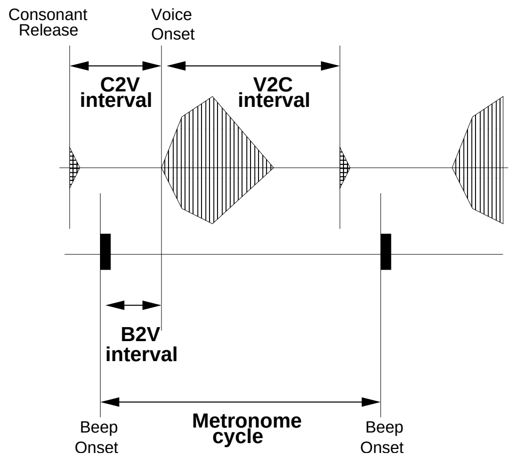

Figure 1: Durational intervals in speech and beep

3 RESULTS

Figure 2 shows a sample data of three intervals (B2V, C2V, V2C) from the voice onset in each cycle. The B2V interval fluctuates around the zero-line more in the beginning and the end of the trial but less in the middle. This suggests a feedback control in a speaker was successfully working in the beginning to the middle of the trial to adjust the time-lag between the voice onset and the beep. As for the latter two intervals, only the shorter one with respect to the vowel in question is plotted. When the metronome rate is slow, V2C is typically shorter than C2V because the speaker starts with a VC syllable. As the rate increases, both C2V and V2C shrink but C2V shrinks rather abruptly when a switch from a VC to a CV syllable occurs. Since V2C is always positive and C2V is always negative, the zero-crossing point of the solid line with dots in Figure 2 corresponds to the VC-to-CV switching point. It is thus quite simple to find the rate-induced resyllabification and its switching point in the plot.

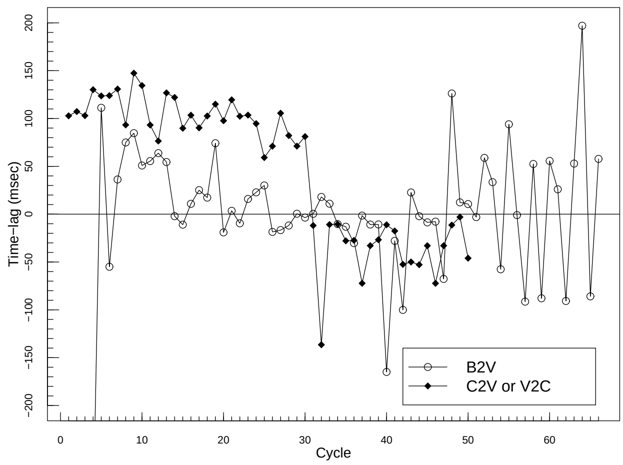

Figure 2: Sample data (Syllable [ib], Subject No.05)

In this particular example in Figure 2, the switch takes place only once in the 30th cycle (between the 30th and the 31st beeps) and there is no ambiguity nor fluctuations over the zero-line. However, there are some cases where the switch is not so obvious. A criterion was set for such cases which requires at least three consecutive positive V2C points followed by at least three consecutive negative C2V points.

As we mentioned in Section 2, frequent epenthesis was a language-particular problem for this paradigm. Some speakers could not resist epenthesizing even if they were told not to do so. In such cases, the expected VC-to-CV switch was hardly observed. Incidentally, some speakers lost the track of the metronome and stop speaking before the end of a trial, which has no switch. More commonly seen was incomplete articulation of segments at the latter half of the trial. Missing or weak burst of plosives, creaky-voiced vowels, incomplete nasals and nasalized vowels were common in most speakers' data. These factors add up to reduce the number of usable data points. Table 1 shows the average VC-to-CV switch point for each consonant along with the number of data points. The maximum number of data points for a segment is 8.

Figure 3 shows an average of switch points for each natural class given in (1)–(3). Relations of natural classes in terms of switch points completely matches those given by sonority scale.

Table 1: VC-to-CV Switching point and the number of data points.

Segments	p	t	k
Mean switch point	8.00	19.80	13.83
# of data points	1	5	6
Segments	b	d	g
Mean switch point	22.25	22.60	25.00
# of data points	4	5	5
Segments	f	s	z
Mean switch point	35.75	38.00	27.50
# of data points	4	3	4
Segments	m	n	l
Mean switch point	30.75	35.00	32.60
# of data points	4	1	5

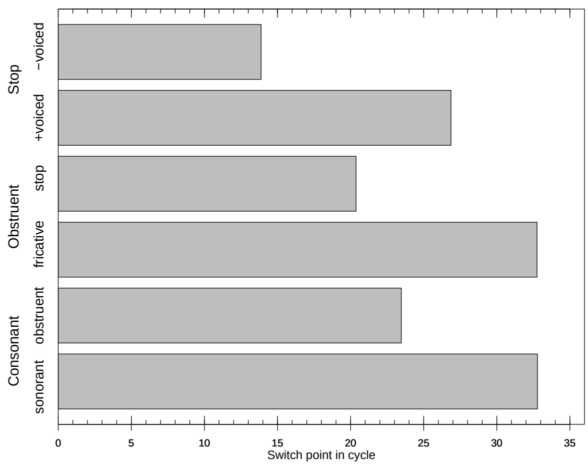

Figure 3: Comparison of switch points across natural classes

4 DISCUSSION & CONCLUSION

Results indicate that (I) most consonants show a clear switch from VC to CV syllable except for [Vn] and [Vp] and (II) the switch point measure closely corresponds to the sonority scale.

As for (I), language-particular factors might account for the exceptional cases. Since [Vn] is the only licit VC syllable word finally in the phonotactics of Japanese, a switch from [Vn] to [nV] is less likely than other VC syllables. [Vp] is the standard material in the rate scaling paradigm in English. However, the phoneme /p/ in Japanese has a quite skewed distribution with respect to vocabulary strata and phonotactics. Word-initial /p/ appears only in mimetics and recent loanwords. Word-internal /p/ does occur but mostly as a geminate in the core vocabulary of Japanese [6].

The results in (II) suggest that segments in the lower position on the sonority scale are more likely to switch earlier in the rate scaling paradigm. Then we have to ask: what does it mean to switch earlier? Our answer is, it corresponds to "more effort", "more difficulty" and "harder to articulate". This interpretation is natural in a model of dynamical systems theory where VC and CV syllables are considered to be attractors in the state space [7]. CV syllables are more stable than VC syllables and act as a stronger attractor. Thus, if the current state is at a VC syllable, and the C in VC needs more effort, we are more likely to be pushed out from the weaker attractor and led to the stronger attractor.

Though the experiment was relatively simple, the switching point measure is an output of a quite complex human system and has no direct correspondence to a physical measure. However, we propose that it is a better way of defining sonority-like scale because it strictly refers to the articulatory domain. There is no need to invoke perceptual notions, such as loudness to explain grammatically controlled syllabification.

ACKNOWLEDGMENTS

We are grateful to Ken de Jong and Robert Port at Indiana University for planting seeds for the core ideas of this project. We also thank students at KIT for participation to experiments. This work is supported by Japan Science Promotion Society (JSPS) Grant-in-aid #13610662.

REFERENCES

[1] Juliette Blevins, "Syllable in phonological theory," in *The Handbook of Phonological Theory*, John A. Goldsmith, Ed., pp. 206–244. Blackwell, Cambridge, MA, 1995.

[2] R. H. Stetson, *Motor Phonetics, 2nd Edition*, North-Holland, Amsterdam, 1951.

[3] Betty Tuller and J. A. S. Kelso, "The production and perception of syllable structure," *Journal of Speech and Hearing Research*, vol. 34, pp. 501–508, 1991.

[4] Kenneth de Jong, "Rate-induced resyllabification reivisited," *Language and Speech*, vol. 44, pp. 197–216, 2001.

[5] Kenneth de Jong, "Effects of syllable affiliation and consonant voicing on temporal adjustment in a repetitive speech-production task," *Journal of Speech, Language, and Hearing Research*, vol. 44, pp. 826–840, 2001.

[6] Timothy J. Vance, *An Introduction to Japanese Phonology*, State University of New York Press, Albany, NY, 1987.

[7] Scott J. A. Kelso, *Dynamic Patterns: The Self-Organization of Brain and Behavior*, MIT Press, Cambridge, MA, 1995.