OCP Effects on Japanese Phonological Processing

Katsuo Tamaoka (Hiroshima University, Japan)
Tadao Murata (Kyushu Institute of Technology, Japan)

Key Words: OCP, Vowel Repetition, Cascaded Dual-route Model, 'Whack-a-mole' Phenomenon

The Obligatory Contour Principle (OCP) refers to a linguistic constraint on similar or same phonological elements or features from being repeated (e.g., Fukazawa, 1999, 2000; Itô & Mester, 1986; Kubozono, 1999; Kubozono & Ôta, 1998).   Kubozono and Ôta (1998) suggested the possibility that vowel dissimilation in Japanese may be a result of the OCP.   For example, the two Japanese morphemes /nana/ meaning 'seven' and /ka/ meaning 'day' combine to form the compound word /nanoka/ meaning 'the seventh day' instead of /nanaka/, which would seem to be the likely combination.   This process of vowel dissimilation occurs so as to avoid vowel repetition of /a/ in sequence within the three mora CVCVCV word structure (C referring to 'consonant' and V referring to 'vowel').   Thus, it would be expected that naming visually-presented Japanese words and nonwords violating the OCP (i.e., same vowel repetition in a series of CV strings) would result in slower processing speeds and higher error rates.   To test this hypothesis, two experiments were conducted. The first experiment examined OCP effects in naming real words with and without same vowel repetition.   The second experiment investigated OCP effects in naming nonwords under both conditions.

EXPERIMENT 1
NAMING REAL WORDS WITH SAME VOWEL REPETITION

Assuming that the OCP affects phonological processing of words with same vowel repetition, the first experiment tested the prediction that CVCVCV-structured words with same vowel repetition would take longer to be named and result in higher error rates than a counter group of CVCVCV-structured words with no such repetition.

Method

Subjects   Twelve female and six male students (18 in total) at Hiroshima University in Japan participated in Experiment 1.   Average age was 27 years and six months for females and 28 years and 3 months for males.   All participants were native speakers of Japanese.

Stimuli   In order to examine OCP effects, words with same vowel repetition were taken from a Japanese dictionary comprised of 50,000 words (Yamada & Ishiwata, 1983).   As shown in Table 1, a total of 175 words with a CVCVCV structure and same vowel repetition such as /sakana/ ('fish'), /hibiki/ ('echo') and /kokoro/ ('heart') were found.   Among the five vowels in Japanese, the vowel /a/ was the most frequently-repeated, found in 62 words (58 nouns, 2 adverbs and 2 adjectives).   The vowel /u/ came in second, repeated in 59 words (11 nouns and 48 verbs).   Since the dictionary form of Japanese verbs always ends with the inflection /u/, it came as no surprise to find so many verbs with the vowel /u/ repeated.   The vowel /i/ was the third most repeated vowel, found in 39 words (only nouns) while the vowel /o/ came in fourth, repeated in 15 words (13 nouns and 2 adverbs).   The

vowel /e/ was only seen repeated in the adverb /semete/ ('at least') out of all 50,000 words in the dictionary.   Kubozono (1998, 1999) explained that the vowels /a/, /i/ and /u/ occur with the greatest distance apart from each other within the 'vowel space', or space in the mouth where vowels are pronounced.   Kubozono (1999) further explained that these three vowels are most frequently found within the various languages of the world.   It is therefore reasonable to expect that the vowels of /a/, /i/ and /u/ are found to be repeated more often in a single word in Japanese than the vowels of /e/ and /o/.

TABLE 1
Real Words with a CVCVCV String
and Same Vowel Repetition

Type of Vowel	Grammatical Categories				Total
	Noun	Verb	Adverb	Adjective	
/a/	58	0	2	2	62
/e/	0	0	1	0	1
/i/	39	0	0	0	39
/o/	13	0	2	0	15
/u/	11	48	0	0	59
Total	121	48	5	2	175

Note: Words above were taken from among the 50,000 words
contained in the *Practical Japanese Dictionary* by Yamada
and Ishiwata (1983).

Fifty-four words of a three mora CVCVCV pattern were used as stimuli and divided up into eighteen sets, each containing three words.   Each of the three words in each set fell under one of three conditions.   The first condition was real words with same vowel repetition as in   /karada/ meaning 'body'.   The second condition was real words with no repetition of same vowels as in

/kurasi/ meaning 'life'.   Word frequencies using the index established by the National Language Research Institute (1973) were matched across the real words in each set under these two conditions. Real words with same vowel repetition (M=20.56) did not differ from real words with no such repetition (M=20.83).   As for the third condition, nonwords without same vowel repetition were created.   The words across all three stimulus conditions in each set began with the same consonant so as to avoid a difference in initial sounds which might have affected naming latency related to the triggering of the voice-activated key.   All stimuli were presented to subjects in the hiragana script.

Procedure   Stimuli of real words and nonwords were individually presented to subjects in the center of a Toshiba J-3100 plasma display at a comfortable distance away in a dimly lit, quiet room.   Stimuli were randomly presented 600 milliseconds after the appearance of an eye fixation point marked by an asterisk '*'.   Subjects were required to pronounce each string of CVCVCV morae as quickly and as accurately as possible.   Subjects spoke into a table microphone connected to a computer via a voice-activated relay mechanism which automatically turned off a timer used to measure naming latency.   Pronunciation was judged as being correct or not by the examiner and then entered into the computer.   The next fixation point was indicated 600 ms after the examiner pressed the space key. Twenty-four practice trials were given to subjects prior to the commencement of the actual testing to familiarize them with the task.

Results

The mean naming latencies and error rates across the three stimulus groups are presented in Table

2. Before the analysis was performed, reaction times outside of 2.5 standard deviations in both the high and low range were replaced by the boundaries indicated by 2.5 standard deviations from the individual means of subjects.   This procedure affected 33 scores (3.4%) including both correct and incorrect responses for stimulus items.   A one-way analysis of variance (ANOVA) with repeated measures was conducted on the reaction times of correct responses and error rates across all three stimulus conditions.   Statistical tests followed analyses of both subject (Fs) and item (Fi) variability.

TABLE 2
Mean Naming Latencies and Percentage of Errors in the Naming Task of Experiment 1

Measure	Stimulus Condition		
	Real Words with Same Vowel Repetition	Real Words with No Repetition	Nonwords
Naming Latency	501 ms (64)	500 ms (66)	570 ms (110)
Error Rates	2.47%	1.23%	8.64%

Note. Figures in parentheses indicate standard deviations.

Naming Latency   Naming latency was based on correct responses from the groups of real words and nonwords.   A one-way ANOVA revealed that there was a significant main effect of stimulus conditions in subject analysis [$Fs(2, 34)=24.76, p<.0001$] and item analysis [$Fi(2, 34)=19.48, p<.0001$]. Further analysis with an orthogonal polynominal comparison was carried out to identify the difference between the three stimulus conditions.   As indicated in Table 2, there was only a 1 ms difference in response time between the naming of real words with and without same vowel repetition.   This difference was not significant [$Fs(1, 17)=0.03, p<.86; Fi(1, 17)=0.00, p<.98$].   Thus, the repetition of same vowels did not affect the naming latency of real words with a CVCVCV structure.   In contrast, significant differences were found between the naming latencies of real words with same vowel repetition and nonwords [$Fs(1, 17)=27.57, p<.0001; Fi(1, 17)=22.50, p<.0005$], and between those of real words without same vowel repetition and nonwords [$Fs(1, 17)=26.48, p<.0001; Fi(1, 17)=22.52, p<.0005$].   Thus, an effect of lexical status was apparent.

Error Rates   The same one-way ANOVA used for naming latencies was carried out on error rates. There was a significant main effect of stimulus conditions in both subject analysis [$Fs(2, 34)=6.56, p<.005$] and item analysis [$Fi(2, 34)=14.07, p<.0001$].   Further analysis with an orthogonal polynominal comparison was carried out to identify the difference in error rates across three stimulus groups.   As indicated in Table 2, the difference was only 1.24 % between real words with and without same vowel repetition.   This difference was not significant [$Fs(1, 17)=1.66, p<.23; Fi(1, 17)=1.56, p<.23$].   Thus, repetition of the same vowel in real words did not affect error rates.   In contrast, significant differences were found between real words with same vowel repetition and nonwords [$Fs(1, 17)=6.12, p<.05; Fi(1, 17)=12.20, p<.005$], and between real words without same vowel repetition and nonwords [$Fs(1, 17)=7.56, p<.05; Fi(1, 17)=21.24, p<.0005$].   Again, an effect of lexical status was apparent from these results.

Discussion

Experiment 1 used 18 real words with same vowel repetition taken from among the 121 CVCVCV- structured nouns found in the Japanese dictionary (Yamada & Ishiwata, 1983) as shown in Table 1.   These 18 real words with same vowel repetition were matched with 18 real words with no such repetition according to phonological structure, number of morae and word frequency.   The results of Experiment 1 indicated no difference in the means of naming latencies and error rates for

pronouncing either of these groups of real words. The OCP showed no effect in the naming of real words with a CVCVCV structure and same vowel repetition. Both conditions of real words with the same repeated and non-repeated vowels resulted in faster naming latencies and lesser error rates than did the condition of nonwords. A lexical status effect was clearly observed in Experiment 1. Consequently, as proposed by the cascaded dual-route model in Figure 1 (Coltheart, 1985; Coltheart & Rastle, K., 1994; Coltheart, Curtis, Atkins & Haller, 1993), it is possible that 'addressed phonology' (the processing of a word as a whole unit) is involved in the processing of real words regardless of whether they have same vowel repetition or not, whereas 'assembled phonology' (the processing of a nonwords by putting smaller phonological units together) is involved in the processing of nonwords. Thus, Experiment 2 investigated the phonological processing of nonwords with a CVCVCV string of same vowels.

EXPERIMENT 2
NAMING NONWORDS WITH SAME VOWEL REPETITION

The naming task in Experiment 1 did not find effects of same vowel repetition on the phonological processing of real words. However, a lexical status effect was evident by the difference in naming latencies between real words and nonwords. To further investigate the effects of same vowel repetition, Experiment 2 produced two types of nonwords: one where the vowel remained constant throughout the nonword (e.g., /ta sa ma ka/) and another where the vowels did not retain their consistency (e.g., /te su mo ki/). If the OCP does in fact influence phonological processing, the first type of nonwords with repeated vowels would take longer to be named and cause higher error rates than the second type with differing vowels.

Method

Subjects Twenty-two female and twenty male students (42 in total) at Matsuyama University in Japan participated in Experiment 2. Average age was 21 years for females and 21 years and 6 months for males. All participants were native speakers of Japanese.

TABLE 3
Examples of Stimuli Used in Experiment 2

Stimulus Conditions	Number of Morae		
	Morae	Morae	Morae
Nonwords with Same Vowel Repetition	/koho/ /muru/	/nohoro/ /reneke/	/hisimini/ /sanataha/
Nonwords without Same Vowel Repetition	/koha/ /muro/	/nohure/ /reniko/	/hisemuno/ /saniteho/
Real Words	/kozi/ /muki/	/noudo/ /reiki/	/hiramaku/ /sakumotu/

Stimuli Words constructed from two to four morae were used as stimuli. They were divided into sets according to their number of morae and placed under three conditions (see Table 3). Each set, therefore, contained three words under three different conditions. All words in each set had the same initial mora. The first condition was nonwords with same vowel repetition as in /ko ho mo/. The second condition was nonwords without same vowel repetition but containing the same consonants as nonwords with same vowel repetition. For example, the stimulus nonword of /ko ha mi/ was constructed using the three consonants of /k/, /h/ and /m/ taken from the aforementioned nonword of /ko ho mo/ where the vowel /o/ was repeated. The vowel /o/ of the first mora was kept

and the second and the third vowels were replaced by /a/ and /i/ respectively. The third condition was real words with the same initial mora as nonwords such as /ko ri tu/ to ascertain the well-established principle of lexical status effect which suggests that real words are pronounced more quickly than nonwords. According to this procedure, thirty-six sets of two-mora words, thirty-six sets of three-mora words and thirty-six sets of four-mora words under all three above conditions were created; a total of 108 sets or 324 stimuli were used.

Stimuli were divided into three counterbalanced lists with an equal number of nonwords (with and without same vowel repetition) and real words. The 42 subjects were also divided into three groups of 14. In this manner, assignment of the three stimulus words in each set was given to a different subject. For example, /ko ho mo/, /ko ha mi/ and /ko ri tu/ (all three words in the same set) appeared separately on three different lists. Thus, no subject saw any more than one of these three stimuli during the same task. This method of assigning items to lists avoided repetition of either rime or onset-nucleus patterns by subjects.

Procedure   The naming task of Experiment 2 was conducted in the same way as Experiment 1.

Results

The mean naming latencies and error rates for the words in each set are presented in Table 4. Before performing an analysis, naming latencies outside of 2.5 standard deviations in both the high and low range were replaced by the boundaries indicated by 2.5 standard deviations from the individual means of subjects. This procedure affected 141 scores (3.1%) including both correct and incorrect responses to stimulus items. A two-way 3 X 3 ANOVA with repeated measures was carried out on naming latencies and error rates relating to the two factors of stimulus conditions and mora structure. Statistical tests follow analyses of both subject (Fs) and item (Fi) variability.

TABLE 4
Mean Naming Latencies and Percentage of Errors
for the Naming Task in Experiment 2

Stimulus Conditions	Number of Morae		
	Morae	Morae	Morae
Nonwords with Same Vowel Repetition	574 ms (115) 6.94%	702 ms (160) 13.69%	804 ms (172) 20.83%
Nonwords without Same Vowel Repetition	545 ms (90) 2.38%	663 ms (139) 6.75%	781 ms (169) 12.10%
Real Words	513 ms (73) 1.39%	586 ms (115) 1.79%	580 ms (108) 1.59%

Note Figures in parentheses indicate standard deviations.

Naming latency   Only correct responses of nonwords and real words were used as data in the analysis of naming latency. The two-way 3 X 3 ANOVA indicated that a main effect of stimulus conditions was significant in subject analysis [$Fs(2, 82)=135.12$, $p<.0001$] and item analysis [$Fi(2, 70)=253.73$, $p<.0001$]. A main effect of the three types of mora structure was also significant in subject analysis [$Fs(2, 82)=186.20$, $p<.0001$] and item analysis [$Fi(2, 70)=176.08$, $p<.0001$]. The interaction of both main effects was also significant [$Fs(4, 164)=60.37$, $p<.0001$; $Fi(4, 140)=38.45$, $p<.0001$]. Real words were named consistently faster than nonwords, from 513 ms for those with two morae to 580 ms for those with four morae. The average naming latency for all real words was 560 ms. The fact that this figure was 133 ms. less than the average for nonwords with the same vowel repeated and 103 ms. less than that for nonwords with no repetition of same vowels confirmed a lexical status effect. Further analysis with an orthogonal polynominal comparison was carried out to identify the difference

between the stimulus conditions of repeated and non-repeated vowels. Significant differences were found with respect to the two stimulus conditions of repeated and non-repeated vowels on two mora [$Fs(1, 41)=16.00$, $p<.0005$; $Fi(1, 35)=10.63$, $p<.005$], three mora [$Fs(1, 41)=15.37$, $p<.0005$; $Fi(1, 35)=10.63$, $p<.005$] and four mora [$Fs(1, 41)=5.00$, $p<.05$; $Fi(1, 35)=6.66$, $p<.05$] nonwords. These results indicate that it took longer to name a nonword when the same vowel was repeated.

Error rates The same ANOVA analysis was conducted on error rates. A main effect of stimulus conditions was significant in subject analysis [$Fs(2, 82)=77.68$, $p<.0001$] and item analysis [$Fi(2, 70)=43.28$, $p<.0001$]. A main effect of the three types of mora structure was also significant in subject analysis [$Fs(2, 82)=32.38$, $p<.0001$] and item analysis [$Fi(2, 70)=16.32$, $p<.0001$]. The interaction of both main effects was also significant [$Fs(4, 164)=9.74$, $p<.0001$; $Fi(4, 140)=6.11$, $p<.0001$]. The average error rate in naming real words was only 1.59%, whereas the average for nonwords was 13.82% with same vowel repetition and 7.08% with no such repetition. The error rates also supported the presumed effects of lexical status. Further analysis with an orthogonal polynominal comparison was carried out. Significant differences were found between the two stimulus conditions of repeated and non-repeated vowels on two mora [$Fs(1, 41)=16.00$, $p<.0005$; $Fi(1, 35)=7.13$, $p<.05$], three mora [$Fs(1, 41)=15.37$, $p<.0005$; $Fi(1, 35)=8.23$, $p<.01$] and four mora [$Fs(1, 41)=5.00$, $p<.05$; $Fi(1, 35)=8.03$, $p<.01$] nonwords. The results of these error rates also confirm effects of same vowel repetition on nonwords.

Discussion

As in the first experiment, effects of lexical status were also indicated in Experiment 2, which is shown by the fact that real words were named more quickly than both types of nonwords. The naming latencies for nonwords with repeated and non-repeated vowels showed that nonwords which had the same vowel repeated were named more slowly than nonwords which did not have the same vowel repeated. In addition, error rates were higher in naming nonwords with repeated vowels than for the non-repeated condition. Both naming latencies and error rates supported OCP effects. Even though nonwords were presented in kana which represent phonological units of mora, it is a smaller unit of phonemes (i.e., vowels in Experiment 2) which affected the assembling process of nonwords.

GENERAL DISCUSSION

The present study sought out to reveal how the OCP concerning same vowel repetition affects phonological processing of Japanese real words and nonwords. The findings can be summarized in three ways. First, real words proved to be named more quickly and accurately than nonwords in both Experiments 1 and 2. In other words, a 'lexical status effect' (real word superiority over nonwords) was clearly observed in both experiments. These results suggest that real words may be processed differently from nonwords. Second, in Experiment 1, real words with same vowel repetition were processed as fast and as accurately as real words with no such repetition. Thus, as far as the chosen 18 stimuli are concerned (about 15% of the 121 total nouns with a CVCVCV string and same vowel repetition found in the dictionary used), the OCP effect does not seem to apply to same vowel repetition in the naming of real words. Third, nonwords with same vowel repetition were named slower and less accurately than nonwords with no such repetition in Experiment 2. This trend was observed in all sets of nonwords containing 2-4 morae in a CV string. Consequently, an effect of the OCP related to same vowel repetition is apparent in the phonological processing of nonwords. The

lexical status effect found in Experiments 1 and 2 provides fundamental evidence to support the cascaded dual-route model as shown in Figure 1 (Coltheart, 1985, 1987; Coltheart & Rastle, K., 1994; Coltheart, Curtis, Atkins & Haller, 1993). In addition to a lexical status effect, the present study showed a clear distinction between real words and nonwords with regards to the OCP. This contrast of OCP effects on the phonological processing of real words and nonwords is nicely explained by the cascaded dual-route model. The phonological processing of real words as a whole unit (i.e., addressed phonology) occurs regardless of same vowel repetition, whereas the processing of nonwords involves the putting together of smaller phonological units (assembled phonology).
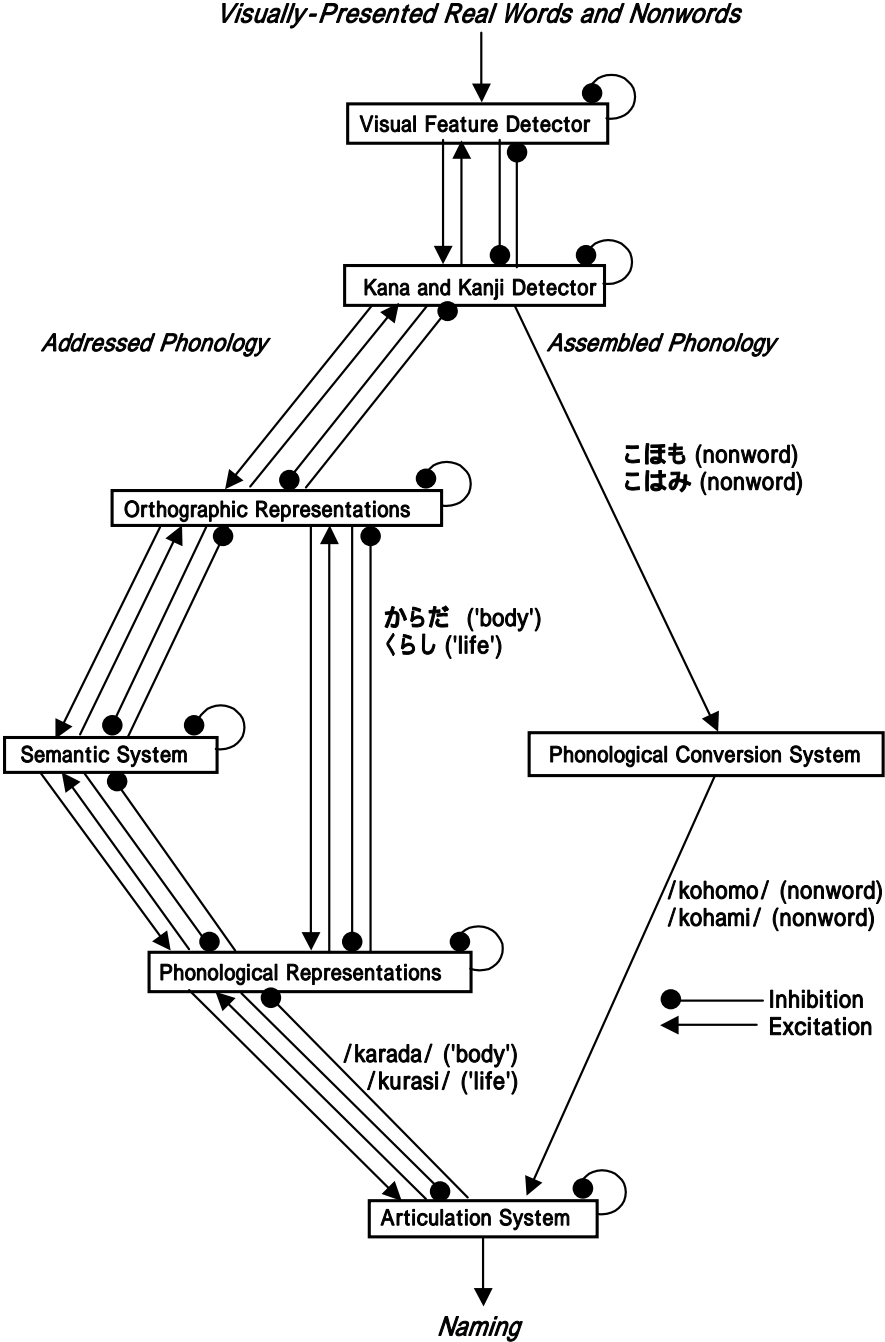

FIG 1. The dual-route cascaded model applied to the Japanese phonological processing of real words and nonwords with and without same vowel repetition

125

The explanation proposed for the putting together of smaller phonological units (i.e., assembled phonology) when processing nonwords is the 'whack-a-mole' phenomenon. The vowel in the first CV mora continues to have a high activation level even when following CV morae are activated. When the same vowel is repeated throughout the CV morae, as in some of the nonwords used in this study, all the CV morae are simultaneously excited to reach the activation level. For example, since the three morae of /kohomo/ share the same vowel of /o/, the first vowel /o/ will be activated to combine with the first consonant /k/ to form the first mora /ko/. However, since the vowel /o/ must also be combined with the second and third consonants /h/ and /m/, it creates a high activation levels of the subsequent morae /ho/ and /mo/. To avoid confusing the sequential order of morae in a CV string, subsequent morae must be inhibited so as not to be activated to the same degree as the previous CV mora. This pattern of activation (or excitation) and inhibition results in the decreased speed of phonological processing of nonwords with same vowel repetition and in the increased rate of pronunciation errors. In contrary, as for naming nonwords with no repeated vowels, phonological assembly of morae in sequential order within a CV string is not affected by the activation of other morae. For example, the nonword /kohami/ does not share the same vowel throughout the CV string so assembling its sequential order of phonological structure is not affected by same vowel repetition. Thus, nonwords with varying vowels are named more quickly than nonwords with repeated vowels and, concomitantly, fewer errors are observed.

REFERENCES

Coltheart, M. (1985). Cognitive neuropsychology and the study of reading. In M. I. Posner & O. S. M. Marin (Eds.), *Attention and performance XI*, 3-37. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Coltheart, M. (1987). Functional architecture of the language-processing system. In M.Coltheart, G. Sartori & R. Job (Eds.), *The cognitive neuropsychology of language*, 1-25. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Coltheart, M., & Rastle, K. (1994). Serial processing in reading aloud: Evidence for dual-route models of reading. *Journal of Experimental Psychology: Human Perception and Performance*, **20**, 1197-1211.

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed processing approaches. *Psychological Review*, **100**, 589-608.

Fukazawa, H. (1999). Theoretical implications of OCP effects on features in optimality theory. Unpublished Ph. D. dissertation submitted to the University of Maryland, USA.

Fukazawa, H. (2000). Typology of OCP on features. *Phonological Studies*, **3**, 121-134.

Itô, J., & Mester, R.-A. (1986). The phonology of voicing in Japanese: Theoretical consequences for morphological accessibility. *Linguistic Inquiry*, **17**, 49-73.

Kubozono, H. (1998). *Onseigaku on'inron [Phonetics and phonology]*. Tôkyo: Kuroshio Shuppan. (in Japanese)

Kubozono, H. (1999). *Nihongo no onsei [Japanese phonetics]*. Tôkyo: Kenkyûsha Shuppan. (in Japanese)

Kubozono, H., & Ôta, S. (1998). *On'in kôzô to akusento [Phonological structure and accents]*. Tôkyo: Kenkyûsha Shuppan. (in Japanese)

Yamada, T., & Ishiwata, T. (1983). Jitsuyô kokugo jiten [*A practical Japanese dictionary*]. Tôkyo: Kadokawa Shoten. (in Japanese)