Proceedings of Meetings on Acoustics

Volume 19, 2013 http://acousticalsociety.org/

ICA 2013 Montreal
Montreal, Canada
2 - 7 June 2013

Speech Communication
Session 5aSCb: Production and Perception II: The Speech Segment (Poster Session)

5aSCb46. Pronunciation of German suffixes by Japanese native speakers of different proficiency levels

Marino Kasuya* and Takayuki Arai

***Corresponding author's address: Sophia University, 7-1 Kioi-cho, Chiyoda-ku, 102-8554, Tokyo, Japan, mar3576ino_7@hotmail.co.jp**

This study investigates an aspect of speech rhythm in German spoken by Japanese native speakers of different proficiency levels. Previous studies on the production of vowel reduction have indicated that this is an area of difficulty for non-native speakers. One study, working on the assumption that second language (L2) speech production is affected by first language (L1), suggested that Japanese native speakers tend to fail at producing the required vowel reductions in unstressed syllables. The present study further investigated this issue by dividing Japanese native speakers into two groups: advanced and elementary learners. The aim of the present study was to investigate acoustic properties of vowel quality (first and second formants) and quantity (durational ratio) of unstressed syllables in German suffixes on the basis of German proficiency levels and the influence of L1. From results, main effect was obtained for the proficiency levels; acoustic analysis showed significant differences between first two formants and durational ratio same to be the factors that caused the difference among the levels. This suggests that L2 learning process may accompany the acquisition of L2 sounds even when rhythmic structures differ between L1 and L2. [Work supported by JSPS.]

Published by the Acoustical Society of America through the American Institute of Physics

©2013 Acoustical Society of America [DOI: 10.1121/1.4801058]
Received 8 Jan 2013; published 2 Jun 2013

INTRODUCTION

Languages have various differences, one of which is their speech rhythm structure. In traditional descriptions of speech rhythm, languages have been classified as "stress-timed" (e.g., German and English), in which stressed syllables occur at equal intervals, "syllable-timed" (e.g., French and Spanish), in which syllables occur at equal intervals and "mora-timed" (e.g., Japanese), in which moras occur at equal intervals. Language speech rhythm is more apt to be influenced by the first language (L1) and vowel reduction is subject to this. Therefore, it is often claimed that foreign language learners generally, and also Japanese native speakers especially, have difficulty in producing their target language.

This paper focuses on vowel reduction of German suffixes, which we consider to be a phenomenon of undershoot related to unattained targets, generally involving shortened vowel duration leading to a reduced vowel space. In the vowels in unstressed syllables, the shortened syllables are not produced canonically, and segments are lenited. However, most previous studies analyzed vowel reduction in English, which is a stress-timed language. Vowel reduction also applies to other languages, but it is slightly different depending on each language, even through some languages have the same category of language speech rhythm.

Previous studies of stress-timed languages suggested that vowels are gradually reduced, depending on speaking rate and individual speaker differences, such as speech style. In particular, Kohler[1] found that vowels in German are more often reduced in function words than content words and more common in spontaneous speech than in passage-reading style and showed that the inflected definite article 'dem' undergoes the following progressively stronger reductions. From left to right, this shows a vowel whose quality or length changes into a schwa. Finally, both vowel and consonant are deleted: de:m → dem → dəm → dm → bm → m.

Also, Kohler[2] suggested that vowel reduction was affected by the place where it occured in the words. All post-stress unstressed syllables are reduced to a schwa in German. In particular, the German suffix C (consonant) + <-en> syllables (i.e. syllables consisting of any consonants plus the <-en> morpheme), which occurs frequently in verb endings (e.g. laufen (to run)), is a typical example in which vowels are reduced in unstressed positions. In connected speech, these vowels may even be deleted. In a corpus analysis of German speech, Helgason & Kohler[3] found that 59% of all vowels in post-stress syllables were deleted, such as in C + <-en>. This phenomenon of vowel reduction has been suggested to be a difficulty facing second language (L2) learners. Yet, only few studies so far have investigated vowel reduction by non-native speakers of German. Kaltenbacher[4] analyzed vowel reduction processes in words spoken in isolation and embedded in short phrases. On the basis of the assumption that native language structures would become evident in L2 productions, she asked native speakers of English, Russian, and Japanese to imitate and read a list of sentences read by a native speaker of German and then analyzed a number of target words. As expected, English speakers reduced vowels in unacceptable places for German speakers, mainly in word-initial unstressed syllables, and produced reduced vowels in function words not compatible with German phonology. The Russian native speakers, conversely, tended to reduce vowels in word-final syllables. The Japanese native speakers, finally, failed to produce the required vowel reductions in unstressed syllables.

Gut[5] suggested that non-native speakers of German produce, in the same passage with the previous research[4], an overall lower amount of reduced vowels than native speakers. She also analyzed the duration of full and reduced vowels produced by non-native speakers of German with three different native languages. She calculated the syllable ratio by dividing the length of each full-vowel syllable by the length of a following syllable with a reduced or deleted vowel and averaging the sum of all ratios by their total number. This measurement indicates the durational difference between syllable pairs consisting of syllables with full vowels and syllables with reduced or deleted vowels, thus reflecting the effect of vowel reduction. For German native speakers, full-vowel syllables are on average 1.87 times longer than syllables with reduced or deleted vowels. All non-native speakers of German, in contrast, show a significantly lower ratio between the duration of non-reduced and adjacent reduced syllables. Syllables with full vowels are on average only 1.5 times longer than syllables with reduced vowels. This suggests that reduced vowels produced by non-native speakers are longer than those produced by native speakers. Comparing the durational extent of vowel reduction among non-native speakers with different L1s, Gut[5] suggests that the lack of durational difference between full-vowel and reduced syllables in non-native speakers of German can be traced back to structural differences between the speakers' L1 and L2. She further found that the vowel reduction of learners of German, measured with the syllable ratio, was more similar to native speech rhythm in the retellings of a story than in the readings of the story.

On the whole, these previous studies suggest that vowel reduction by non-native speakers of German, especially Japanese native speakers[4], differs significantly from that in German native speakers. Yet, a more systematic

approach is necessary to capture the nature of the deviation from native speech. The previous studies analyzed non-native speech auditorily[4] or measured the duration of syllables containing unstressed vowels[5]. Unstressed vowels have not yet been acoustically analyzed enough in terms of their duration and formant structure. Moreover, although most studies analyze vowel reduction in English in the stress-timed language, only few studies so far have investigated vowel reduction by non-native speakers. Furthermore, it cannot be decided yet whether vowel reduction is influenced by the L1 as claimed by Kaltenbacher[4] and Gut[5], or whether it constitutes a universal difficulty for language learners. Therefore, this study investigates vowel reduction in German spoken by Japanese native speakers of different proficiency levels.

The present study has two aims:

- to analyze acoustic properties of unstressed syllables of German suffix by Japanese native speakers both in terms of duration and vowel formant structure
- to investigate of the effect of different German proficiency levels of non-native speakers

EXPERIMENT

Participants

The Participants were ten Japanese native speakers with two different German proficiency levels (five advanced and five elementary learners), and five native speakers of standard German.

Materials and analysis

The participants were instructed to utter at different speaking rates: slow, normal, and fast. The speech materials consisted of word stems containing /_Cen/ sequence, where C stands for any consonant and the vowels are realized as schwa in native German, as for example in geben (to give), reden (to read), and legen (to put). The participants were asked to read the target words embedded in the carrier phrase 'Ich have /_Cen/ gesagt.' ('I said /_Cen/.'). All ten sentences were repeated three times, and 1350 tokens (15 participants *10 sentences * 3 speaking rates * 3 times) were analyzed.

To analyze the acoustic properties of vowel formant (F1/F2) and durational ratio using Praat[6], the duration of all vowels and consonants in the German suffix C (consonant) + <-en> syllables was measured using a phonetic criteria relying on Gut[7] and Peterson & Lehiste[8] (at the beginning of a stable formant structure, especially at the onset of the first formant, and end at the end of a stable formant structure, especially at the end of the second formant). The frequencies of the first and second formants (F1/F2) were measured at each vowel midpoint.

Results

This section describes the results classified into two cases: the results of measurement of vowel quality depending on the formant values (F1/F2) and vowel quantity depending on the durational ratio of German suffixes. The unstressed vowels were categorized by the degree of vowel reduction, depending on the previous study, which suggested that unstressed vowels undergo progressively stronger reduction[1].

- non-reduced vowel (a full vowel): /e/, /ɛ/
- reduced vowel: /ə/
- deleted vowel (no waveform nor formants)

First, the following shows the sound quality change of the vowels in German suffixes. To measure of unstressed vowel quality, this study intends to investigate "non-reduced vowel" and "reduced vowel". Therefore, the production data for "deleted vowel" are not plotted in figures 1-3. Figures 1 through 3 depict how F1/F2 values of reduced vowels are distributed in the speakers' F1/F2 formant space. In the native speakers' speech data, as predicted, vowel quality tend to be reduced towards schwa. Also, there was a strong contrast between advanced learners and elementary learners.

Figure 1 illustrates the vowel quality change in different speaking rates of German native speakers. They produced reduced vowels with a mean F1 of 503 Hz (S.D. = 143 Hz). For F2, reduced vowels showed a mean F2 of 1693 Hz (S.D. = 349 Hz). However, they also produced deleted vowels, so, this figure shows only part of the reduced ones.

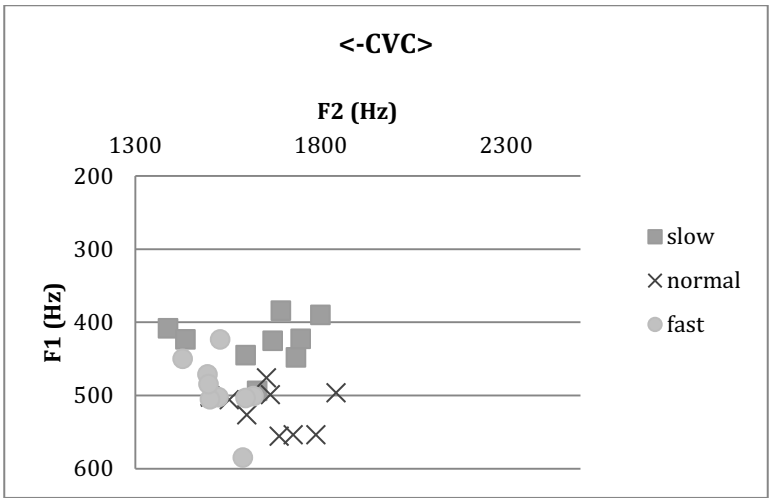

FIGURE 1. Vowel formant values F1/F2 of German suffixes <-CVC> by German native speakers

Figure 2 illustrates the vowel quality change in different speaking rates of advanced learners. They produced reduced vowels with a mean F1 of 498 Hz (S.D. = 156 Hz). For F2, reduced vowels showed a mean F2 of 1705 Hz (S.D. = 338Hz). They produced different vowel quality at each speaking rate. Although advanced learners tended to produce unstressed vowels clearly at slowly speaking rate, Figure 2 indicated that the vowel quality of them changed to schwa by increasing their speaking rates.

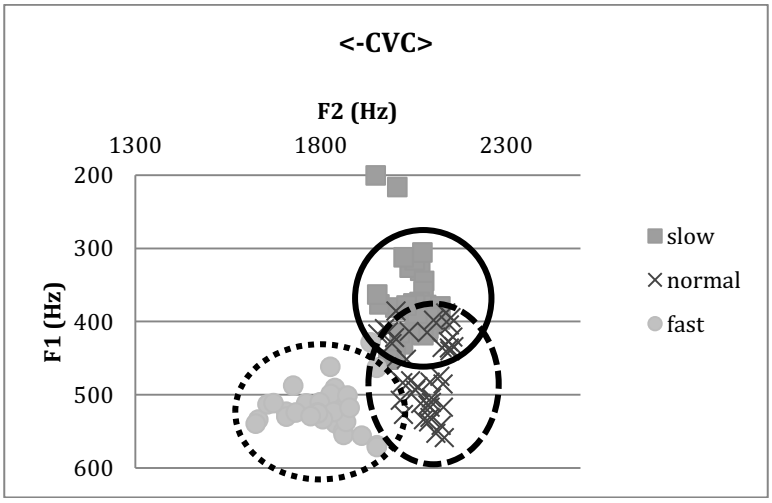

FIGURE 2. Vowel formant values F1/F2 of German suffixes <-CVC> by advanced learners

Figure 3 illustrates the vowel quality change in different speaking rates of elementary learners. They produced non-reduced vowels with a mean F1 of 467 Hz (S.D. = 152 Hz). For F2, non-reduced vowels showed a mean F2 of 2158 Hz (S.D. = 402 Hz). They did not show any systematic formant shifts due to the speaking rate.

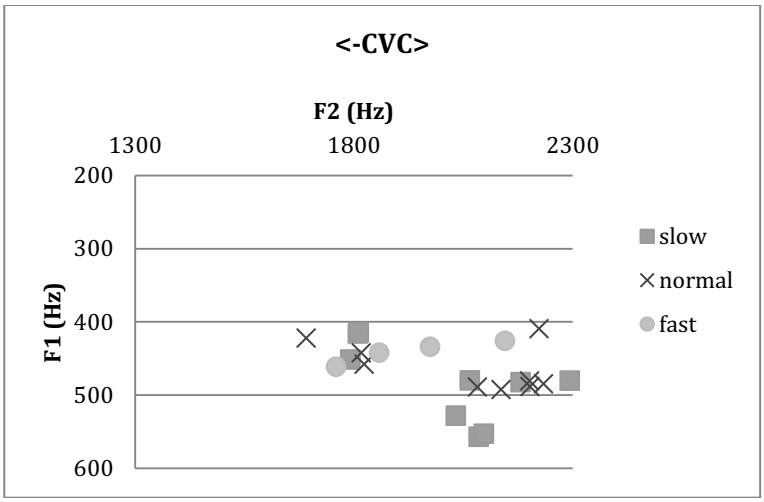

FIGURE 3. Vowel formant values F1/F2 of German suffixes <-CVC> by elementary learners

There was clear spectral difference between advanced learners and elementary learners. Advanced learners appear to have made a distinction. The dispersion from the centroid of the F1/F2 formant space was measured. The dispersion values were subjected to repeated measurements by ANOVA with two independent variables: the subject group (advanced learners, elementary learners, and German native speakers as a control group) and speaking rate (slow, normal, and fast), and the dependent variable, "dispersion from centroid." The analysis revealed the main effect of the subject group ($F=41.616$, $p<.001$): the quality of reduced vowel was changed by their speaking rate and by increasing proficiency levels of German. In particular, the results of post-hoc analysis indicated elementary learners and advanced learners significantly differed ($p<.0001$), although advanced learners and German native speakers did not. Also, elementary learners and advanced learners significantly differed in terms of each speaking rate.

German native speakers and Japanese German learners also had clear durational difference. By changing the speaking rate, German native speakers produced shorter vowel duration /e/ and longer duration of ending of words /n/. On the other hand, Japanese German learners produced relatively sufficient vowel duration /e/ even when changing the speaking rate. In particular, the durational ratio of advanced learners was similar to that of native speakers. On the other hand, elementary learners produced continuously sufficient vowel duration /e/ even if their speaking rate was increasing.

The following shows the change of the durational ratio in German suffixes. To measure of durational ratio of unstressed syllables, this study intends to investigate "non-reduced vowel", "reduced vowel" and also "deleted vowel". Therefore, all production data are plotted in figures 4-5. Figure 4 illustrates the mean of durational ratio of the German suffix <-CVC> in different speaking rates of German native speakers.

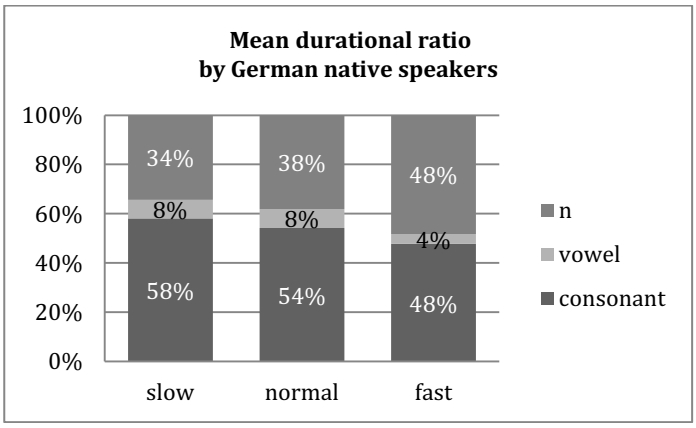

FIGURE 4. Mean durational ratio of German suffixes <-CVC> by German native speakers

Figure 5 illustrates the mean of vowel durational ratio of the German suffix <-CVC> in different speaking rates of advanced learners.

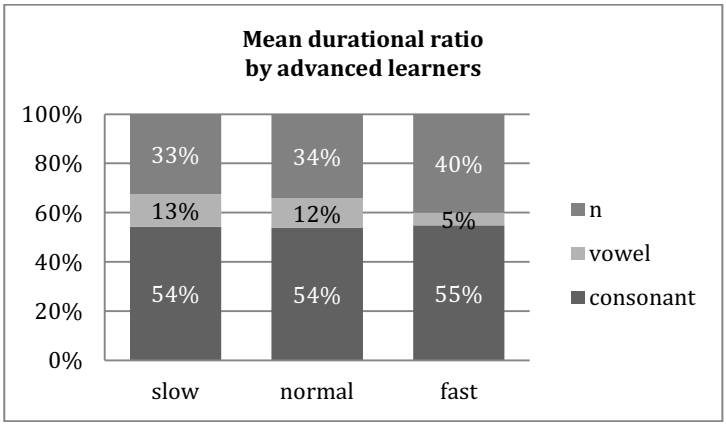

FIGURE 5. Mean durational ratio of German suffixes <-CVC> by advanced learners

Figure 6 illustrates the mean of vowel durational ratio in different speaking rates of elementary learners.

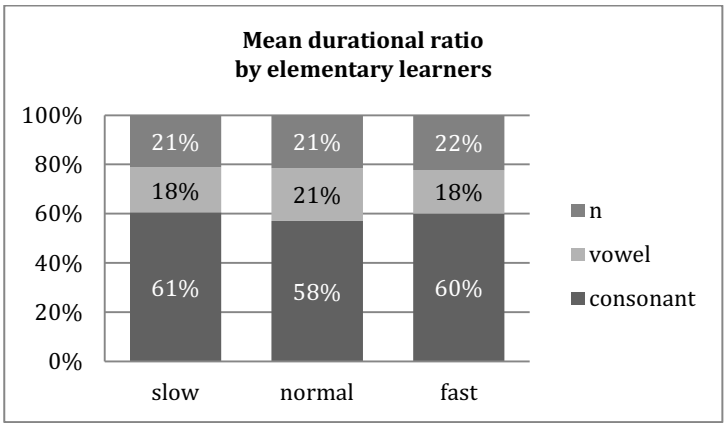

FIGURE 6. Mean durational ratio of German suffixes <-CVC> by elementary learners

Advanced and elementary learners also displayed clear durational difference. The durational ratio of the reduced vowel in unstressed syllables, which consisted of consonant and vowel in the German suffix C + <en> syllables, was measured and their mean values was measured by each subject group. The values were subjected to repeated measurements by ANOVA with two independent variables: the subject group (advanced learners, elementary learners, and German native speakers as a control group) and speaking rate (slow, normal, and fast), and the dependent variable, "durational ratio of reduced vowel." The analysis revealed the main effect of the subject group **(F=**82.323, $p<.005$): the duration of reduced vowel was changed by their speaking rate and by increasing proficiency levels of German. In particular, the results of post-hoc analysis indicated that elementary learners and advanced learners significantly differed ($p<.0001$), although advanced learners and German native speakers did not. Also, elementary learners and advanced learners significantly differed in terms of each speaking rate.

The statistical results indicated that durational ratio tended to affect also non-native speech. There was clear significant difference by increasing their proficiency levels in both vowel quality and durational ratio in the utterance. Although some previous studies indicated that reducing vowels in unstressed syllables was difficult for Japanese native speakers, this study provided that this can be overcome by increasing proficiency levels, that is their learning experience of German.

DISCUSSION

This study investigated vowel reduction, an aspect of speech rhythm of German, spoken by Japanese native speakers of different proficiency levels and had the following two aims:

- to analyze of the acoustic properties of unstressed syllables of German suffix by Japanese native speakers both in terms of duration and vowel formant structure
- to investigate of the effect of different German proficiency levels of non-native speakers

As shown in the production test, unstressed syllables of German suffixes were changed at each speaking rate, and as learners proficiency levels increased, changes tended to come in vowel quality and quantity, namely, vowel formant values F1/F2 and durational ratio of /-CVC/ sequences.

According to previous studies, vowels in /-CVC/ sequence in German suffixes are often deleted the same way they are removed by German native speakers. Japanese German learners tended to produce fast definitely non-reduced vowels: [e] and [ɛ]. However, this research clear contrasts with previous research. Vowels in /-CVC/ sequence in German suffixes "e" in advanced learner production are relatively close to the mid vowel [ə], which is called "schwa", in the native vowel space. However, what elementary learners produced was clearly close to non-reduced vowels, in particular, [e] and [ɛ] in the native vowel space, as previous studies indicated. That is, their production of German [ə] carried the L1 Japanese color of [e] and [ɛ]. The tongue position in their productions of [ə] showed lower F1 and higher F2 than that of native speakers and advanced learners.

Moreover, participant groups clearly differed in terms of the durational ratio of suffix /-CVC/. Japanese German learners have relatively sufficient vowel duration regardless of speaking rates. In particular, elementary learners produced the longest vowel duration of three speaker groups. However, most vowel duration produced by advanced learners was similar to that by German native speakers.

The following tables list the overall percentages of the vowels produced by each participant group, which were categorized as either

- non-reduced vowel (a full vowel): /e/, /ɛ/
- reduced vowel: /ə/
- deleted vowel (no waveform nor formants)

TABLE 1. Overall percentages of non-reduced, reduced and deleted vowel syllables produced by German native speakers at each speaking rate (mean values).

Vowel type	Slow	Normal	Fast
Non-reduced vowel	0 (0%)	0 (0%)	0 (0%)
Reduced vowel	401 (89%)	216 (48%)	77 (17%)
Deleted vowel	49 (11%)	234 (52%)	373 (83%)

TABLE 2. Overall percentages of non-reduced, reduced and deleted vowel syllables produced by advanced learners at each speaking rate (mean values).

Vowel type	Slow	Normal	Fast
Non-reduced vowel	333 (74%)	162 (36%)	9 (2%)
Reduced vowel	117 (26%)	275 (61%)	324 (72%)
Deleted vowel	0 (0%)	13 (3%)	117 (26%)

TABLE 3. Overall percentages of non-reduced, reduced and deleted vowel syllables produced by elementary learners at each speaking rate (mean values).

Vowel type	Slow	Normal	Fast
Non-reduced vowel	450 (100%)	450 (100%)	414 (92%)
Reduced vowel	0 (0%)	0 (0%)	36 (8%)
Deleted vowel	0 (0%)	0 (0%)	0 (0%)

As a result, German native speakers produced reduced and deleted vowels in the unaccented vowels, as previous studies suggested. Gut [5] suggested that the Japanese native speakers finally failed to reduce vowels in unstressed syllables in particular, on the durational ratio. However, this study showed both vowel quality and quantity, that is formant and durational ratio, changed progressively increasing proficiency levels. Namely, their production tends to be similar to that of German native speakers. In addition, their percentage was increased by speaking rates: Most advanced learners tended to produce progressively "non-reduced vowel (/e/ and /ɛ/)", "reduced vowel (/ə/)" and only a few "deleted vowels (incl. syllabic nasal)". Elementary learners produced mostly "non-reduced vowel (/e/ and /ɛ/)".

However, there are points, which should be considered in the future research. From the results, vowel articulation might to be nevertheless difficult for Japanese German learners to learn. That is, there were characteristic differences in the changing F1/F2 formant values. Regarding this point, it is assumed that the age at which one stays in Germany for the first time was also related. Most advanced learners in this study were students of the Department of German Studies at Sophia University who had stayed for at least one to three years in Germany. The utterance of the advanced learners slightly differed depending on the age at which they started learning German (They can be divided into two types of groups: those who stayed in German for the first time under the age of twelve, and those who did so later). Their utterances who started learning German before twelve years old seemed to be more similar to those of the native speakers' than those of the learners who started learning German after twelve years old. That is, the former's productions showed changes in both F1/F2 values more similar to those by native speakers than the later. Also, they tended to produce more remarkable utterances at F1, which is concerned with the degree of mouth opening at the lips. After that, the F2 (the position of articulation) showed changes in the values. However, there could not be shown differences statistically yet, so this research should be analyzed further by increasing the number of participants.

ACKNOWLEDGMENTS

I would like to give heartful thanks to Akiko MASAKI whose help were of inestimable value for my study. I would also like to thank anonymous reviewers whose opinions and information have helped me very much throughout the production of this study. This work was supported by JSPS KAKENHI Grant Number 2410524.

REFERENCES

[1]K. Kohler, "Segmental reduction in connected speech in German: phonological facts and phonetic explanations." In W.J. Hardcastle & A. Marchal (ed.), Speech Production and Speech Modeling. Dordrecht: Kluwer Academic Publishers (1990).

[2]K. Kohler, "Variability of opening and closing gestures in speech communication", In Sound Patterns in German Read and Spontaneous Speech; Symbolic Structures and Gestural Dynamics. K. Kohler (ed.). Arbeitsberichte des Instituts für Phonetik und digitale Sprachverarbeitung Universität Kiel 35, 33-96 (2001).

[3]K. Helgason and K. Kohler, "Vowel deletion in the Kiel Corpus of Spontaneous Speech", In Sound Patterns in Spontaneous Speech, K. Kohler; C. Rehor; A. Simpson (ed.). Arbeitsberichte des Instituts für Phonetik und digitale Sprachverarbeitung Universität Kiel 30, 115-157 (1996).

[4]E. Kaltenbacher, "Zum Sprachrhythmus des Deutschen und seinem Erwerb", In Eine zweite Sprache lernen, H. Wegener (ed.). Tübingen: Narr, 21-38 (1998).

[5]U. Gut, "Non-native speech rhythm in German", Proceedings of the ICPhS, Barcelona, 2437-2440 (2003).

[6]P. Boersma and D. Weenink, "Praat, a system for doing phonetics by computer", Glot International 5:9/10, 341-345 (2001).

[7]U. Gut, "Unstressed vowels in non-native German", Proceedings of the 3rd ICSP, Dresden (2006).

[8]G. Peterson and I. Lehiste, "Duration of syllable nuclei in English", J. Acoust. Soc. Am 32, 693-703 (1960).