Ann.Bull.RILP(1987)No.21, 147-151

PRELIMINARY REPORT ON THE EFFECTS OF SPEAKING RATE UPON FORMANT TRAJECTORIES

Satoshi Imaizumi, Shigeru Kiritani, Hajime Hirose,
Shin-ichi Togami* and Katsuhiko Shirai*

Introduction

Although there have been several studies on the effects of speaking rates on formant trajectories[1-11], rules for speech synthesis taking into account these effects are still under investigation.

There are still numerous differences among studies on the effects of speaking rate. Some papers[1,2,11] indicate that increased rates of speech result in systematic deviation of the obtained formant values from their putative targets, that is, "vowel reduction". Some others[3,4] insist that such "vowel reduction" does not always occur at faster speaking rates. Furthermore, still others[8,9,10] claim that adjustments in speaking rate are achieved by strategy which differ among speakers.

The differences mentioned above may be due to differences in the speech material being analyzed, including segmental-suprasegmental features, context and manner of speaking of individuals. Much more precise analyses for speech data together with data concerning the articulatory dynamics, collected under various conditions from a number of speakers seem to be necessary to construct synthesis rules describing how speaking rates affect formant trajectories.

This paper reports some preliminary results of our experiments to understand the effect of speaking rates on formant trajectories and to construct synthesis rules for natural sounding speech.

Method

The speech material analyzed were the Japanese vowels symbolized by V_2 in nonsense words of the form /$V_1CV_2CV_1$/, where V_1 and V_2 were either /a,i/ and C was one of /b,d,g/. All possible combinations resulted in 12 nonsense words. The phoneme /d/ was uttered as /dz/ when it preceded /i/. Each of the 12 words was embedded in a carrier sentence / korewa --- desu/, which means /this is ---/, and was repeated five times in random order.

The subject was an adult, male, native speaker of the Tokyo dialect of Japanese. Three speaking rates were studied: a slow, normal and fast rate. Each speaking rate was based on the subject's own appraisal of his comfortable rates. The subject was asked to speak naturally and clearly.

*Waseda University

Using the ultrasonic technique reported by Niimi and et al.[12], the tongue configuration in the midsagittal and frontal planes during utterances was observed and recorded on a video tape recorder. The EGG (Electro-glottogram) signal was also recorded simultaneously for future use. The speech and EGG waveforms were recorded using a PCM recorder.

The formant trajectories were extracted from the LPC analysis based on the autocorrelation method.

Results and Discussion

Figure 1 shows the first and second formant trajectories for some of the utterances spoken by the subject. Because the Japanese phoneme /d/ is usually uttered as /dz/ when it precedes /i/, the utterances /VdVdV/ were omitted from this figure. In Fig. 1, each curve indicates the average formant trajectory for five /VCVCV/ utterances. Even for the stop consonants, the first and second poles were extracted from the LPC analysis. These pole frequencies were connected smoothly to construct the first and the second formant trajectories. The solid lines indicate the results for the slow speaking rate, and the dotted lines that for the fast speaking rate. In Fig. 1, several differences in the formant trajectories can be observed between the two rates of speaking.

First, comparing the middle vowels in /aCaCa/ and /iCiCi/, for instance, it was found that only the effect of speaking rate on the first formant trajectory was different. For the middle /a/ in both /ababa/ and /agaga/, the average trajectories of the first formant at the fast speaking rate were smaller than those at the slow speaking rate. The second formant of the middle /a/ in these utterances, on the other hand, did not change much with variation in the speaking rate. For the middle /i/ in /iCiCi/, both the first and the second formant trajectories remained unchanged even though the speaking rate varied. For the middle vowels in the /aCaCa/ and /iCiCi/ sequences, it seems that the "undershoot" effect at the fast speaking rate occurred only for the first formant of /a/.

Next, comparing the middle vowels in /aCiCa/ and /iCaCi/, both the first and the second formant trajectories of /a/ were found to have larger variations with the speaking rate than those for /i/. The first formant trajectories of the middle /i/ remained almost unchanged for both /ibabi/ and /igagi/ when the speaking rate increased. Thus, for the middle vowels in the /iCaCi/ and /aCiCa/ sequences, the "undershoot" effect at the fast speaking rate seemed to occur for /a/ but not for /i/.

Furthermore, in Fig. 1, the effects of speaking rate on the formant trajectories for the middle /a/ seem different between /aCaCa/ and /iCaCi/. For the middle vowel /a/ in /aCaCa/, only the first formant trajectories vary with the speaking rate. On the other hand, for the middle /a/ in /iCaCi/, both the first and

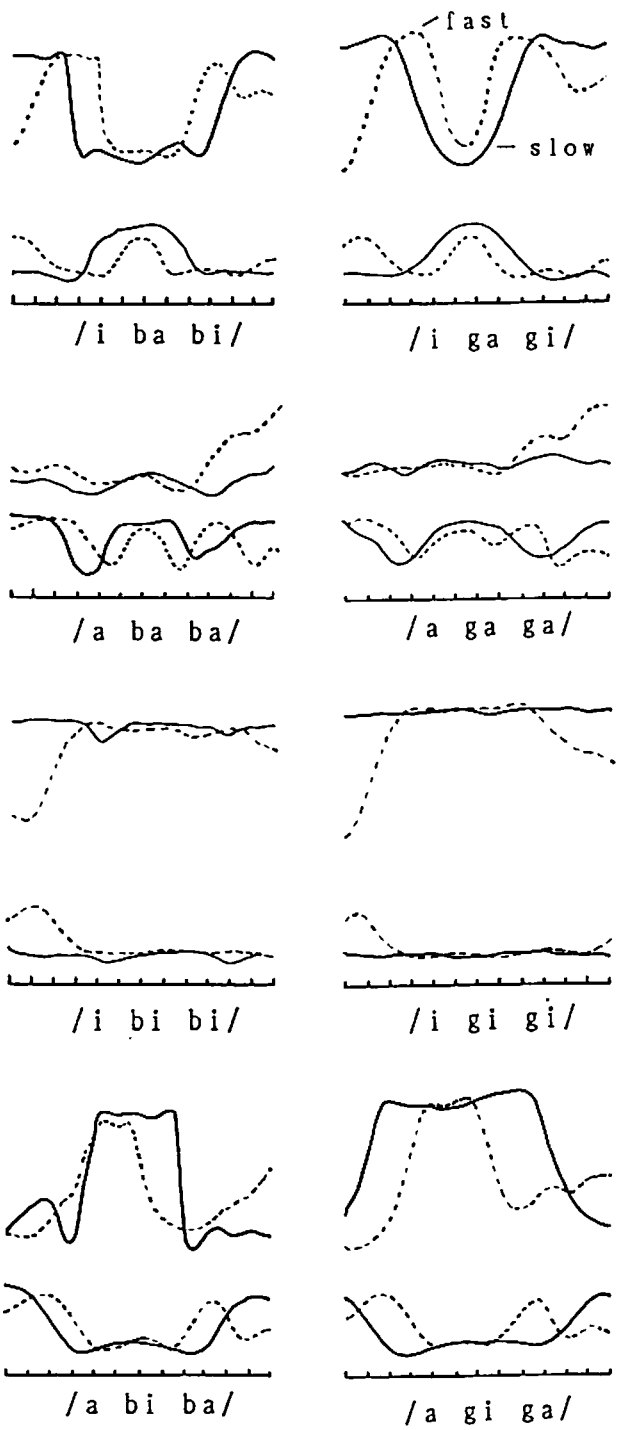

Fig. 1. The effects of speaking rate and context on first and second formant trajectories for the Japanese vowels /a/ and /i/ in several contexts.

the second formant trajectories vary largely when the speaking rate increases. This may indicate that the effects of speaking rate on the formant trajectories of the middle /a/ in /aCaCa/ and /iCaCi/ are different depending upon the vowels preceding or following the /CaC/ portion.

These results indicate that the effects of speaking rate on formant trajectories are different between /a/ and /i/. For all contexts discussed in this paper, the vowel /a/ showed greater variation in both the first and the second formant trajectories than /i/ when the speaking rate increased. Also, the variation in the middle /a/ in /CaC/ were affected by the vowels preceding or following the sequence /CaC/. This may mean that the effects of speaking rate are different depending not only upon the vowels analyzed, but also upon their preceding or following vowel-consonant sequences.

Much more intensive analyses of speech data, and also data concerning the articulatory dynamics, collected under various conditions from a number of speakers are necessary to draw more detailed conclusions.

Summary

This paper describes some preliminary results of our study on the effects of speaking rate upon the formant trajectories of Japanese vowels. The results show that the effects of speaking rate are different depending not only upon the vowels analyzed, but also upon their preceding or following vowel-consonant sequences.

Acknowledgment

This work was supported in part by Grant-in-Aid for Scientific Research on Priority Areas, "Advanced Man-Machine Interface Through Spoken Language", the Ministry of Education, Science and Culture, Japan.

References

1) B. Lindblom: Spectrographic study of vowel reduction. J. Acoust. Soc. Am., 35(11),1773-1781, 1963.
2) T. Gay: Effect of speaking rate on diphthong formant movements. J. Acoust. Soc. Am., 44, 1570-1573, 1968.
3) R. R. Rerbrugge and D. Shankweiler: Prosodic information for vowel identity. J. Acoust. Soc. Am., 61, S39, 1977.
4) T. Gay: Effect of speaking rate on vowel formant movements. J. Acoust. Soc. Am., 63(1), 223-230, 1978.
5) D. O'Shaughnessy: The effects of speaking rate on formant transitions in French synthesis-by-rule. Proc. 1986 IEEE-IECEJ-ASJ, Tokyo, 2027-2030, 1986.
6) J. Miller and T. Baer: Some effects of speaking rate on the production of /b/ and /w/. J. Acoust. Soc. Am., 73(5), 1751-

1755, 1983.
7) J. L. Miller: Effects of speaking rate on segmental distinctions. in Perspectives on the study of speech. P.D. Eimas and J.L. Miller (Eds.), Lawrence Erlbaum Associates, New Jersey, 39-74, 1981.
8) D. P. Kuehn and K. L. Moll: A cineradiographic study of VC and CV articulatory velocities. J. Phonetics, 4, 303-320, 1976.
9) Y. Sonoda and K. Nakakido: Effect of speaking rate on jaw movements in vowel sequences. J. Acoust. Soc. Jpn.(E), 7, 1, 5-12, 1986.
10) Y. Sonoda: Effect of speaking rate on articulatory dynamics and motor event. J. Phonetics, 15, 145-156, 1987.
11) S. Nakajima: Rules of vowel formant modification in accordance with speech rate. Trans. IEICE Committee on Speech Research, 17-24, 1986.
12) S. Niimi, S. Kiritani and H. Hirose: Ultrasonic observation of the tongue with reference to palatal configuration. Ann. Bull. RILP, 19, 21-27, 1985.