The Timing Control of Speech Production Adapted to Delayed Auditory Feedback

SAWADA Haruhiko[†] and KAKEHI Kazuhiko[†‡]
†Graduate School of Human Informatics, Nagoya University, Japan
‡Center for Integrated Acoustic Information Research (CIAIR), Nagoya University, Japan
E-mail: sawada@cog.human.nagoya-u.ac.jp

ABSTRACT

The purpose of this study is to examine the effect of delayed auditory feedback on reaction time and duration of speech. Six normal subjects participated in this experiment and were exposed to delayed auditory feedback (delay times were 0ms, 100ms, 200ms, 300ms or 400ms). After reading a simple text during exposure to delayed auditory feedback, the duration of the speech and the reaction time in the delayed naming task were measured, using a CRT tachistoscope. Subjects showed not only a significant prolongation of duration of speech, but also a significant delay of reaction time for words, especially on the 200ms condition of delayed auditory feedback. This indicates that the speech planning process might predict and the delay of auditory feedback before articulation. Furthermore, word size effect in reaction time was observed independent of delayed auditory feedback.

1. INTRODUCTION

One interest of speech production study concerns speech monitoring. Speech monitoring is the process by which speakers are able to check the correctness and appropriateness of their speech. The most widely accepted monitoring theory has been proposed by Levelt (1989).

Levelt (1989) proposed four production stages and one feedback stage of speech. Production stages are composed of lemma selection, phonological encoding, phonetic encoding and articulation. Feedback stage consists of two pathways of monitoring. One involves perception of self-produced overt speech and the other involves internal speech. However, Levelt's model was set up to clarify the speech production system based on segmental structure (conceptual, lexical, phonological, or phonetic). The covert monitoring has been discussed and detailed by many researchers including himself. The overt monitor about segmental information also has been done well. But how to monitor temporal information has not been sufficiently investigated.

The main purpose of this study is to reveal the overt monitoring system, which detects the errors of temporal information and repairs the timing of speech. For this, it is important to examine which production stage (conceptual, lexical, phonological, phonetic, or as a whole) should process the temporal information of overt auditory feedback.

We investigated the effect of delayed auditory feedback, which influences the speech monitoring system. The study of delayed auditory feedback began in the 1950s by Lee (1950). In the study, it was proved that a 200ms delay time of auditory feedback influences articulation, e.g., prolonging articulation and increasing speech error and loudness. This means that auditory information of self-speech should be fed back to some production stages of speech, which repairs the speech plan. But there are no experimental studies that show how the delayed auditory feedback influences the timing control of speech production. How the monitoring system detects the delay of auditory feedback and repairs the timing control of speech production needs to be examined.

To reveal the timing control of speech production via a monitoring system, we measured reaction time and duration of speech during the course of a delayed naming task. Duration of speech indicates prolongation of articulation, which easily shows the effect of delayed auditory feedback on the monitoring system. Analysis of the reaction time shows differences of the temporal process of speech production from the experimental conditions. This is a paradigm which has been used in previous speech production studies. Sternberg et al. (1978) showed a word size effect on reaction time which increases linearly with the number of words. Reaction time provides critical information inferring to the planning process of speech production. By using the delayed naming task, the planning process of speech production could be separated from the perception or recognition stages. According to Levelt's model, the planning process with the delayed naming task might be the stage concerned with phonetic encoding.

In this report, the timing control of speech production was examined using a delayed naming task.

2. EXPERIMENT

To examine the effect of delayed auditory feedback on the speech monitoring system, we conducted a delayed naming task.

ISBN 1-876346-48-5

Subjects were exposed to delayed auditory feedback for 3 hours as a whole. Delay times of auditory feedback were changed to 0ms, 100ms, 200ms, 300ms and 400ms. At each condition, the subject carried out the delayed naming task after reading a simple text under exposure of delayed auditory feedback. In this task, reaction time and duration of speech were important features to investigate the characteristics of timing control. The experimental conditions were arranged to know the effects of the delay time of auditory feedback and number of moras.

2.1. Method

Subjects

Six Nagoya University undergraduate students participated in this experiment (3 males and 3 females). All subjects were native Japanese speakers and had no history of language disorders. The preparatory experiment confirmed that delayed auditory feedback significantly influenced all subjects who participated in this experiment.

Stimuli

The stimuli for the delayed naming task consisted of two sets of words. These stimuli were selected for variation of word size. One set of words consisted of Japanese 2-mora words “たか” /taka/ and “かた” /kata/. The other set of words consisted of Japanese 6-mora words “たかたかたか” /takatakataka/ and “かたかたかた” /katakatakata/.

Apparatus

The experiment was controlled by an IWATU ISEL AV tachistoscope system (IS 701D). The auditory signals of targets and naming responses of the words were recorded using a SONY Digital Audio Deck (DTC-2000ES) and SONY electret condenser microphone (ECM-530/1). The audio samples were stored for off-line analyses as described below.

To realize delayed auditory feedback, subjects used a Victor Headphone (HP-DX1) connected to a BOSS Super Effects Processor (SE70) and spoke into the microphone connected to the processor. Loudness of auditory feedback was arranged to mask the bone conduction feedback. The experimenter could easily change delay time of auditory feedback by setting the dial of the processor.

Experimental Condition

Conditions of five delay times of auditory feedback (0ms, 100ms, 200ms, 300ms and 400ms) were used as the auditory condition. Delay time of 0ms means normal auditory feedback.

Two sets of words (2 mora and 6 mora) were used as the number of mora conditions in this experiment.

Task

The experiment consisted of two tasks. One was a simple reading task. Subjects just read a simple text under exposure of delayed auditory feedback to measure the effect of delayed auditory feedback.

The other was a delayed naming task. First, a subject pressed the key to start the task and memorized a word presented at the center of the display. Then, a subject pressed the key again and spoke the word as accurately and as soon as possible when a target signal appeared. The target signals were controlled to appear at the center of the display within a few seconds (randomly varying from 2 to 4 seconds) after the key pressing. The target signals for subjects were visual signals (○:white circle as “go” or ×:cross mark as “no go”) and at the same time an auditory signal (800Hz tone-100ms duration) was recorded to be used for measuring the reaction time. A subject then pressed the key to move on to the next trial.

Procedure

The experiment consisted of 21 trial blocks. Each block included 1 trial of a simple reading task and 20 trials of a delayed naming task. The order of 20 trials was randomized.

The first block was used for practice and the 2nd, 7th, 12th and 17th blocks were used for the 0ms condition of delayed auditory feedback. The 3rd to 6th blocks, the 8th to 11th blocks, the 13th to 16th blocks and the 18th to 21st blocks were used for the 100ms, 200ms, 300ms and 400ms conditions of delayed auditory feedback, respectively. The order of the conditions of the delayed auditory feedback was counterbalanced across subjects.

Analysis

In this study, all of the audio samples, audio signal for target and speech response of delayed naming task were recorded onto a digital audiotape. After the experiment, duration of speech and reaction time of each trial were measured in each delayed naming task. Duration of speech was determined as the duration between the onset of speech and the end determined by visual and auditory inspection. Reaction time was measured as the interval between the onset of audio marker and the onset of speech.

3. RESULT

In the experiment of the simple naming task, all subjects were affected by delayed auditory feedback. They prolonged speech, increased errors and spoke more loudly. The results of a simple reading task show the effect of delayed auditory feedback that peaks at about the 200ms condition. The results were consistent with those in past studies as well as in the delayed naming task of the experiment described below.

The experimental results of the delayed naming task were shown in Figure 1 and 2 to depict effects of delayed auditory feedback.

Figure 2 shows the mean duration of speech over all subjects. The left side shows the results of 2-mora words and the right side, 6-mora words. The horizontal axis shows delay time conditions of auditory feedback and the vertical axis shows duration of speech in ms. The difference of the duration of speech between a 2-mora word and a 6-mora word in no delay conditions is due to the difference of mora length of the word. In a 2-mora word case, the duration of speech looks irrelevant to the delay time conditions. However, in a 6-mora word case, it seems that the duration of speech reaches the maximum in the 200ms condition.

The results of ANOVA for duration of speech of 2-mora words did not reveal a significant effect of delayed auditory feedback. On the other hand, the results of ANOVA for duration of speech of 6-mora words revealed a significant effect of delayed auditory feedback ($p<0.01$). A multiple comparison by Tukey`s honestly significant difference (HSD) test showed that duration of speech of 6-mora words increased from the 0ms to 200ms condition and decreased from the 200ms to 400ms condition. Therefore, the 200ms delay of auditory feedback showed the largest effect.

Figure 2 shows the mean reaction time over all subjects. The left side shows the results of 2-mora words and the right side, 6-mora words. The horizontal axis shows the delay time conditions of auditory feedback and the vertical axis shows reaction time in ms. It seems that reaction time of a 6-mora word is longer than that of a 2-mora word in the same delay time condition. This is the word size effect on reaction time. The reaction time in 200ms delay time conditions is the largest both for a 2-mora word and a 6-mora word.

The results of ANOVA for reaction time revealed a significant effect of word size ($p<0.01$) and delayed auditory feedback ($p<0.01$). A multiple comparison by Tukey`s honestly significant difference (HSD) test of auditory conditions showed that reaction time increased from the 0ms to 200ms condition and decreased from the 200ms to 400ms condition. Therefore, a 200ms delay of auditory feedback showed the largest effect on reaction time as well as on duration of speech.

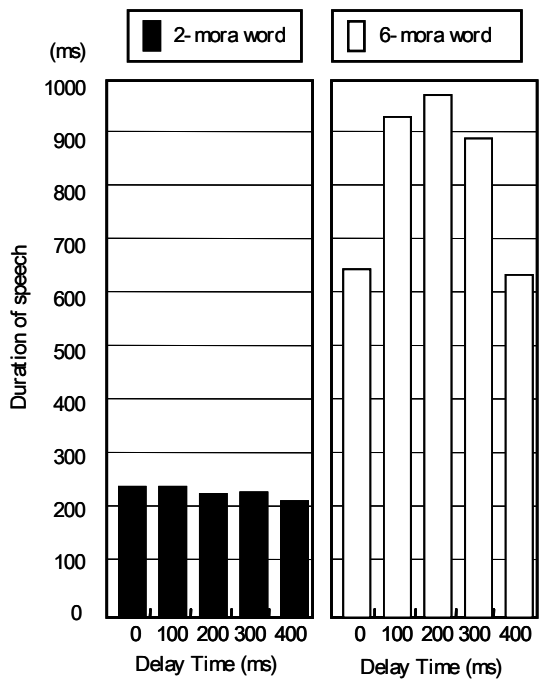

Figure 1. Duration of speech

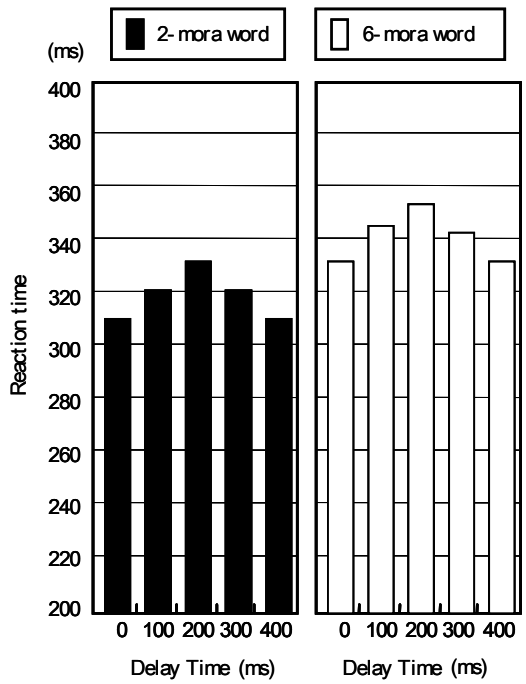

Figure 2. Reaction time

ISBN 1-876346-48-5 © 2003 UAB

4. DISCUSSION

This study was set up to reveal the overt monitoring system, which detects the errors of temporal information and repairs the timing of speech. The data from the experiment showed some critical evidence for the monitoring system and speech timing control.

First, subjects prolonged their articulations of a 6-mora word but not of a 2-mora word. This result is almost consistent with past delayed auditory feedback studies. Delayed auditory feedback interfered with duration of speech of a 6-mora word especially on the condition of a 200ms delay time but not with a 2-mora word, because a 2-mora word could be spoken within a few hundred milliseconds before auditory feedback occurred, while a 6-mora word could not be completed before auditory feedback. This result indicates that the unit size in speech timing control is more than 2 moras or 200ms. In other words, timing control might be conducted on every 2 moras.

The second finding is the most important discovery of this experiment. Subjects delayed their onset of articulation by the effect of delayed auditory feedback. Delayed auditory feedback interfered with reaction time of a 2-mora word and a 6-mora word in the same ways, especially maximized on the condition of 200ms delay time. It is a very interesting new finding related to the monitoring system of temporal information. Although in the course of reaction time, there is no auditory (delayed) feedback, the reaction time is influenced by delay time conditions. It indicates that some kind of adaptation to delay time conditions occurs in the speech production system, which is related to the timing control of speech. It seems that delays of auditory feedback were predicted and influenced on speech planning, even before articulation.

These results about reaction time and duration of speech indicate that the effect of delayed auditory feedback on articulation itself, e.g., prolonging articulation and increasing speech errors and loudness, might be caused by a planning process of timing.

Finally, word size effect was observed as an independent effect from the one of delayed auditory feedback in this experiment. As the reaction time of a 2-mora word increased from the 0ms to 200ms condition and decreased from the 200ms to 400ms condition, so also did the reaction time of a 6-mora word. This result indicates that the number of moras might influence the speech production stages, where a sequence of phonetic codes is prepared, but the delay of auditory feedback might not influence the same stage. It implies that the phonetic encoding and the timing control should be considered as different stages of speech production.

Now, we have a new experimental result about the monitoring system, which influences the timing control of speech production and delays the reaction time when auditory feedback delays for 200ms. But this effect has no relation to the phonetic plan. In Levelt's model, temporal information and segmental information are fed back to higher production stages including the conceptual stage together via a monitoring system. However, even though production stages higher than phonetic encoding are not essential to the delayed naming task in this experiment, temporal information from delayed auditory feedback influenced the results. Furthermore, our result indicates that temporal information did not influence phonetic encoding directly. An alternative root of feedback for temporal information must be supposed.

This result implies that the timing control component of speech production system predicts the delay of auditory feedback by adaptation and affects on the latency utterance. The part of timing control of speech production is supposed independent of the segmental process, like phonetic encoding.

5. CONCLUSIONS

In this study, to revel the monitoring system of overt speech, the effect of delayed auditory feedback on speech timing control was experimentally examined by delayed naming task.

The results show that the effect of delayed auditory feedback on reaction time and duration of speech, and that the effect is independent of word size effect. These results indicate that there might be an adaptive timing control component to auditory feedback in speech production system, which is independent of planning component of segmental codes.

REFERENCES

[1] Lee, B.S., "Effects of delayed speech feedback", Journal of Acoustical Society America, 22:824-825,1950.

[2] Levelt, W. J. M. "Speaking: From intention to articulation". Cambridge, MA: MIT Press.1989.

[3] Sternberg, S. et al. "The latency and duration of rapid movement sequence". Information processing in motor control and learning (Stelmach, G. E), 117-152. New York Academic Press.