Published in final edited form as:

Percept Mot Skills. 2012 December ; 115(3): 845–864.

Effects of delayed auditory feedback on speech kinematics in fluent speakers

Jayanthi Sasisekaran, Ph.D.

University of Minnesota

Summary

The effects of delayed auditory feedback on behavioral responses and speech kinematics were investigated. Participants were 10 males (M age = 22.2, SD = 3.7) and 10 females (M age = 20.5, SD = 2.2). Participants were required to repeat short (six-syllable) and longer (11-syllable) non-words in a reading task. Speech errors, lip movement variability, and movement duration of labial trajectories under delayed feedback were compared to effects under gated speech feedback and a control or synchronous auditory feedback condition. Repeated measures analysis indicated lower percent of correct productions, higher variability and slower duration of lip movements in the delayed feedback condition compared to the gated feedback and control conditions. Implications for auditory feedback of movement control in continuous speech and for theories of delayed auditory feedback that attribute a role to movements are discussed.

Keywords

Delayed auditory feedback; gating; movement variability; movement duration

Delayed auditory feedback is the process of playing back one's own speech after a brief time delay in order to study the role of auditory feedback in the control of speech production. Several studies have been conducted to investigate the effects of delayed feedback on speech production (e.g., Lee, 1950; Goldiamond, 1965; Howell, Powell, & Kahn, 1983). Some of the known effects include decrease in speech rate, increase in speech intensity and pitch, prolongation of vowels, misarticulations, and disfluencies (Goldiamond, 1965; Perkins, Bell, Johnson, & Stochs, 1979; Howell & Archer, 1984). Factors such as age, sex of individual, and task complexity, are known to modulate the influence of delayed feedback (e.g., Siegel, Fehst, Garber, & Pick, 1980; Corey & Cuddapah, 2008). However, there are only a limited number of studies of the effects of delayed feedback on the kinematics of ongoing speech movement (e.g., Mochida, Gomi, & Kashino, 2010; Cai, Gosh, Guenther, & Perkell, 2011). Such effects highlight the role of auditory monitoring of speech during normal speech production and allude to the critical role played by auditory processes in the control of spatial and temporal parameters of coordinated speech production.

Effects of Delayed Auditory Feedback on Speech: Theoretical Considerations

Theories of delayed auditory feedback attribute the effect to different causal mechanisms. For instance, servomotor accounts (e.g., Lee, 1950; Fairbanks, 1955) hypothesize that a comparator looks for a match between intended vocal output and sensory feedback. The

disruptions observed during delayed auditory feedback are thought to be corrective actions to overcome the discrepancies between intended output and the conflicting sensory feedback. Arguments against the servomotor accounts contend that the speech rate is too fast for auditory feedback to be processed and for corrections to be implemented before the subsequent segment (e.g., Borden, 1979). This suggests that anatomical and time constraints may limit any observable effects of delayed auditory feedback on movement kinematics. Yet another account is that delayed feedback creates a disruption in the rhythm of the speech signal by creating an altered rhythm (Disruptive Rhythm Hypothesis, Howell, *et al.*, 1983).

The various theoretical accounts make differing claims about the role of movement. For instance, the EXPLAN theory of production (Howell, 2004) claims that altered auditory feedback can cause disruptions by perturbing the timing of execution. Thus, it suggests that: a) effector movements may be linked to the effects of altered feedback, and b) such effects are not specific to just delayed auditory feedback, but to all forms of altered auditory feedback that disrupt timing. The Node Structure Theory (MacKay, 1987) attributes the observed disruption to changes in the sensitivity of nodes in a neural network that are used for both perception and action and suggests that the effects of delayed auditory feedback are independent of movement. Thus, study of the consequences of delayed auditory feedback on speech movements can explicate the role of auditory feedback in ongoing control of speech and provide a test between theories that postulate a potential link between such effects and movement kinematics. Furthermore, comparing the effects of delayed auditory feedback to other forms of altered feedback can elucidate the underlying mechanisms altered by such processes.

Delayed Auditory Feedback and Movement Kinematics

Studies of delayed auditory feedback on movement kinematics in tasks, such as, rhythmic finger tapping, indicate that the most robust effects include asynchronies between perception and action and increases in timing variability (Pfordresher, 2003; Pfordresher & Benitez, 2007). For instance, Pfordresher and Dalla Bella (2011) tested the hypothesis that disruption depends on the state of the finger movement trajectory at the onset of delayed auditory feedback. Participants tapped isochronous rhythms at a rate specified by a metronome while hearing delayed auditory feedback (for piano tones) of differing lengths. Analyses of finger movements indicated that finger velocity was significantly associated with the effect of delayed auditory feedback on mean inter-response intervals (IRIs) whereas finger position the variability of IRIs. A second experiment supported the conclusion that the effect of delayed auditory feedback on timing variability is strongly associated with the finger trajectory. Such findings argue for paradigms that involve comparison of both temporal and spatial variables in the study of the effects of delayed auditory feedback on movement trajectories.

Delayed auditory feedback and speech kinematics

In fluent speakers, delayed auditory feedback affects motor behavior by disrupting the forward flow of speech. Although the effects on movement control have been investigated, there has been limited study of the direct and immediate effects of auditory alterations on speech movements. Auditory feedback serves as the immediate source for the dynamic control of speech articulation and any disruption to this source of control may result in disruptive speech motor coordination (e.g., Sussman & Smith, 1971; Zimmerman, Brown, Kelso, Hurtig, & Forrest, 1988). Mochida, *et al.* (2010) studied whether temporal changes in auditory feedback of bilabial plosives immediately affect the subsequent lip movement. Participants were required to say the isolated syllable /pa/ seven times while maintaining a constant speech rate. During the repetition, auditory feedback corresponding to the third

repetition was altered by shifting the timing ahead or behind the anticipated production onset and/or replacing the type of syllable while the subsequent feedback was blocked. Labial movement trajectories under altered and normal feedback were compared within a single cycle of lip closing/opening movement subsequent to the auditory alteration. The authors found that a quickening of lip closing/opening was elicited when the auditory feedback preceded the real production by 50 ms. A similar effect was not found when the feedback preceded the real production by more than 50 ms or when it was delayed and when the feedback syllable was replaced by another syllable. The authors interpreted the results as suggesting an underlying mechanism that detects discrepancies between anticipated and actual auditory consequences for the rapid modification of subsequent movements. Cai, *et al.* (2011) manipulated spatial and temporal parameters of formant trajectories and found that compensations to spatial perturbations were bidirectional and opposing the perturbations. Significant adjustments in syllable timing were observed under temporal perturbations as well. Because the above-mentioned studies have used monosyllabic, consonant – vowel combinations as stimuli, the question remains as to what effects will be evident on speech movement coordination in continuous speech under delayed auditory feedback.

There have also been reports of an effect of delayed feedback on the timing of speech movements (Lane & Tranel, 1971; Borden, 1979; Howell & Sackin, 2002). While the most obvious effect is a reduction in speech rate, there have been questions raised about the possibility of disruptions in the timing of the speech movements. Given the high rate of speech sequencing and the relatively long latency of the auditory feedback control loop (120 ms or greater, e.g., Purcell and Munhall, 2006), it remains questionable whether auditory feedback can have an effect on movement timing.

Purposes of the Present Study

Earlier studies of the effects of delayed auditory feedback in fluent speakers have reported an increase in errors and disfluencies (e.g., Perkins, Bell, Johnson, & Stochs, 1979; Howell & Archer, 1984; Corey & Cuddapah, 2008; Lincoln, Packman, Onslow, & Jones, 2010). There have been limited investigations of the effects of delayed auditory feedback on speech movement kinematics (Sussman & Smith, 1971; Mochida, *et al.*, 2010; Cai, *et al.*, 2011). Such investigations are critical as they highlight the role of auditory feedback on the normal process of speech motor control while supporting theoretical accounts of delayed feedback that link it to movement control (e.g., the EXPLAN). In the present study, the link between speech kinematics and the disruptions under delayed auditory feedback was investigated using a paradigm adapted from Smith, Johnson, McGillem, and Goffman (2000). This paradigm offers the advantage of measuring Lip Aperture Variability, a composite measure of spatial and temporal movement variability associated with changes in the size and kinematics of the aperture of the lips for repeated productions. A delay of 200 ms has been reported to have a maximum effect on the speaker as it roughly matches the average syllable length, thus creating an intrusive disruptive rhythm that makes monitoring of the speech signal difficult for the listener (Fairbanks, 1955; Howell, *et al.*, 1983). Therefore, in the present study a delay of 200 ms was introduced to the speech signal. The primary aim was to assess the effects of delayed feedback on the behavioral (percent correct responses) and kinematic responses (movement variability and movement duration of the lip aperture trajectory) in fluent speakers.

The second aim was to compare the effects of delayed feedback with other forms of altered feedback, e.g., gated speech, to investigate if such effects on speech output are unique or if other types of altered feedback have similar consequences. Gating is the process of interrupting on-going speech (by gating it on and off). Earlier studies of gating attempted to

document situations other than delayed feedback which influence speech (Miller & Licklider, 1950; Howell, *et al.*, 1983). Such studies indicated that interrupting speech has similar effects on speech control as delaying it. Therefore, in the present study in addition to delayed auditory feedback, the effect of gating, another type of feedback manipulation, on speech was studied. Based on earlier reports that a gating frequency of 2 Hz (500 ms “on” cycle, 500 ms “off” cycle) is most effective in disrupting speech (Miller & Licklider, 1950; Howell, 2007), this frequency was used to gate speech in the present study. Performances under delayed feedback and gating conditions were compared to a control, synchronous feedback condition.

Earlier studies of delayed auditory feedback have reported some sex-based differences in behavioral (Fukawa, Yoshioka, Ozawa, & Yoshida, 1988; Corey & Cuddapah, 2008) and electrophysiological responses (e.g., Swink & Stuart, 2012). For instance, Corey and Cuddapah (2008) investigated the effects of delayed feedback in the speech of men and women in reading and conversation. They found more stuttering-like-disfluencies in conversation than reading, and a larger effect in men than women. Men also produced more reading errors than women. Such findings suggest that men and women process altered auditory feedback differently while speaking (Swink & Stuart, 2012). Therefore, third aim of the study was to investigate if the earlier reported differences between men and women under delayed auditory feedback are observed in the present study.

Finally, past studies of the effects of delayed auditory feedback on speech kinematics have used monosyllabic, consonant – vowel sequences (Moshida, *et al.*, 2010; Cai, *et al.*, 2011). Such sequences do not allow ready generalization to longer utterances. Therefore, in the present study a nonword reading task was employed to investigate the effects of altered auditory feedback on continuous speech. Nonwords were chosen as they require negligible lexical access and limit the involvement of higher-level language processes (e.g., semantics) while allowing a window into the speech motor output processes. Varying nonword length (six- vs. 11-syllables) allowed the investigation of the effects of task complexity. Past studies have also reported a negligible effect of delayed auditory feedback when a single syllable was repeated, in contrast to the dramatic effect in continuous speech (Correy & Cuddapah, 2008; Chase, 1958; MacKay, 1987). Such findings suggest that longer and more complex utterances will exhibit higher movement variability and movement duration under altered auditory feedback.

Research hypotheses

Using a nonword reading task, four research hypotheses were tested. First, the behavioral (percent errors) and kinematic (movement variability and movement duration of lip aperture trajectories) responses under delayed auditory feedback will differ from those under a control, synchronous feedback condition. Second, behavioral and kinematic responses to delayed feedback will be comparable to such responses under gating. Third, men and women will differ in their behavioral and kinematic responses to delayed auditory feedback. Finally, the magnitude of the observed effects on speech errors and movement kinematics under delayed auditory feedback will be larger for the 11-syllable nonwords compared to the six-syllable nonwords.

METHOD

Participants

Participants were 10 males ($M_{age} = 22.2$, $SD = 3.7$) and 10 females ($M_{age} = 20.5$, $SD = 2.2$), all were monolingual English speakers who passed a hearing screening at 20 dB HL at . 5, 1, 2, 4, and 8 KHz. All participants were undergraduate students recruited using fliers

posted on campus at the University of Minnesota. The experimental protocol was approved by the Institutional Review Board, University of Minnesota.

Participant selection was based on responses to a screening form that was used to rule out positive history of language, speech, hearing, and/or neurological deficits, and current usage of drugs likely to affect the outcome of the experiment (e.g., drugs for ADHD and anti-anxiety drugs). Data from two male participants were excluded from the final data set as 38% and 55% of the total number of trials were error productions in these individuals. Therefore these participants did not fulfill the minimum criteria requirement of four to six correct productions in each of the three conditions (delayed auditory feedback, gating, and control).

Stimuli

Nonwords were used as target stimuli in order to avoid familiarity with the target items. Four nonwords that were six-syllables long and four nonwords that were 11-syllables long formed the target stimuli. All nonwords started with “*mæb*” and ended with the phoneme /b/. This strategy was used to select consistent start and end points for the extraction of oral movement trajectories on the basis of lower lip peak opening velocities for the kinematic analysis. The six-syllable nonwords were: /*mæbbeIbab Ū bibo*/, /*mæbtatabomatiba*/, /*mæbfrŪglabrŪpliba*/, and /*mæbgrafropIŪkriba*/ . The 11-syllable nonwords were /*mæbbeIbabŪbIbotatabomatiba*/, /*mæbbeIbabŪbIbograpropIŪkriba*/, /*mæbfrŪglabrŪplibatabomatiba*/, and /*mæbfrŪglabrŪplibagrafropIŪkriba*/ . All nonwords carried an alternating strong – weak stress pattern. These nonwords were originally constructed to vary in complexity, for instance, /*mæbbeIbabŪbIbo*/ is the least complex nonword while /*mæbgrafropIŪkriba*/ is the most complex nonword in the six-syllable length category as these nonwords contained the least and most complex speech sound combinations. However, the complexity variation within the nonwords was not taken into consideration for analysis purposes.

Apparatus

Delayed feedback and gating protocol—Figure 1 depicts the instrumental set-up of the delayed feedback and gating protocol. The experimental system consisted of: a) a microphone for speech input, b) a pair of insert earphones for presentation of the feedback signal to the participant, and c) a computer-controlled signal processing system. The processor was a Capybara 320 with Kyma X PC software for processor control (both developed by Symbolic Sound Inc.). Kyma X provides modules for various modifications of sound signals. These modules are combined and the resulting code is downloaded to the Capybara 320. The speech input was delayed or gated and had its signal-to-noise ratio (S/N) adjusted by the processor. This signal was fed back to the listener using the earphones at a delay of 200 msec or a gated frequency of 2 Hz. Feedback intensity was pre-set for each participant by calibrating speech output to a comfortable loudness level. Speech intensity was not measured, but observation suggested that speech intensity increased when delayed auditory feedback was presented. Such increases in speech intensity resulted in increases in feedback intensity. The modules in the Kyma X environment included a delay component. In order to calibrate the delay introduced by the processing of the signal through the system, this delay module's parameter was set to 0 ms. The calibration signal was a click-like sound with a sharp leading edge. The input to the system and its output were conveyed to separate channels of a two-channel oscilloscope. The difference between the onset of the processed and unprocessed signals was measured to be 10 ms. This minimal delay was incorporated into all the calculations involving the nominal setting of the Kyma X delay module.

Kinematic procedure—Participants were seated in front of a commercial Optotrak Certus camera (Northern Digital, Waterloo, Ontario). This system allowed tracking of movements in 3D with an accuracy of less than 0.1mm. Eight small (7 mm) infra-red light emitting diodes were attached to the participant's face to track articulatory movements of the upper lip, lower lip, and jaw. Four of the diodes were mounted on a set of goggles that participants were required to wear during the experimental session. One diode was placed in the center of the forehead. Together, these five diodes were used to calculate the 3D head coordinate system (Smith, *et al.*, 2000), which allowed head movement artifact to be eliminated. The remaining three diodes were placed one on the vermillion border of the upper lip, one on the center of the lower lip (this marker represents combined actions of the lower lip and jaw), and one on a splint taped to the jaw. The diodes were fixed to the participant and facing the Optotrak camera such that the movements could be tracked. The diode motions were sampled at a rate of 250 Hz. A condenser wireless microphone was placed approximately 8 cm away from the participant's mouth and a microphone receiver was used to record the acoustic speech signal. This acoustic signal was digitized on an A/D channel of the Optotrak system and was thus synchronized with the movement data. The acoustic signal was digitized at the rate of 16,000 Hz.

Procedure

The experiment consisted of a 1.0 hour session. Two experimenters were present throughout the experimental session and monitored participants' performance. Participants were instructed that they would be provided with eight nonwords in a sheet of paper and they were required to read the nonwords at comfortable loudness. Following instructions, participants were provided two practice trials to ensure that they are able to produce the target nonwords correctly at least once. During practice the eight nonwords were read by the participant twice in a block of eight each. Productions during practice were monitored and corrected for pronunciation and stress errors. The data indicated that a majority of participants were able to achieve a minimum of one correct production within the two practice blocks. Following practice, participants were instructed that the experimenter would call out a number between one and eight and they were required to read the nonword corresponding to the number. They were instructed that they would hear themselves while reading the nonwords, but in some of the trials this feedback will be altered such that it will sound either "delayed" or "choppy". They were asked to try to be as accurate as possible in reading the nonwords.

A total of 18 blocks were presented to all participants. Each block consisted of the eight test nonwords read out aloud once in random order. A number (between one and eight) corresponding to each nonword was presented and participants were required to read out the nonword corresponding to the number. A ten second gap was presented between each nonword production. Of the 18 blocks, six blocks of nonwords were presented with normal feedback (i.e., feedback through earphones placed in the ear), six blocks with delayed feedback, and six blocks with gated feedback. The order of the three types of feedback conditions was randomized across participants. Participants' response during each trial was recorded in order to ensure a minimum four to six correct productions of nonwords in each category at each nonword length as these numbers of trials were minimally required for the kinematic analysis. Up to six additional blocks (two blocks in each condition of delay, gating, and control) were presented in instances where a minimum of four to six correct productions could not be obtained within the first 18 blocks.

The procedure required both the experimenters to record the participants' errors online during the experiment. Post-experiment, one of the experimenter listened independently to the recordings and provided offline scores of correct and error productions for the behavioral

and kinematic analysis. The online coding was later compared to offline analysis of the recorded productions to establish intra- and inter-rater reliability. Only those productions that both raters (offline and online) agreed upon as being correct were included in the kinematic analysis.

Data Analysis

Error analysis—Correct responses were defined as those trials which were free of articulation errors or disfluencies. Both consonant and vowel errors were coded. Articulation errors included substitutions, omissions, distortions, or additions of speech sounds. Disfluencies including, interjections, pauses (> 1 sec), sound or syllable repetitions, prolongations, and blocks, were also coded as errors. Because stress placement tends to vary with delayed auditory feedback, differences in syllable stress were not coded as errors. However, observation indicated that a majority of participants were able to maintain the rhythm of production as the strong syllables in each nonword were underlined to encourage consistent stress placement during reading. This analysis resulted in percent correct productions for each nonword length and condition.

Kinematic analysis—A detailed description of this analysis can be obtained from Smith, *et al.* (2000) and Smith and Zelaznik (2004). The MATLAB digital signal processing software was used to simultaneously load and analysis the acoustic and kinematic data. Following low pass filtering (cut-off 10 Hz) of the lip displacement signals (both upper and lower lip), the velocity signal was computed from the lower lip displacement signal. Following this, the displacement trajectories associated with four to six correct productions of each nonword were obtained by identifying the correct productions from the 18 blocks and segmenting the corresponding displacement file for each correct production. The velocity signal was used to obtain the start and end point of the displacement trajectory of each correct production. The start point for each segment was the point of maximum opening velocity for /m/ (in the first syllable /mæb/) and the end point was the point of maximum opening velocity of the lower lip opening gesture for /b/ (in the last syllable). The peak velocity regions were easily located by the experimenter, and a MATLAB algorithm selected the peak opening velocity within a user-defined window. Each trim corresponding to a correct production was selected in this manner was then reassessed for fluency and accuracy of extraction by listening to the associated audio record.

Following extraction of the lip displacement trajectories for a set of four to six correct productions corresponding to each nonword, a multi-step analysis was performed using a custom MATLAB program. First, the lip aperture (LA) signal was obtained by a sample-by-sample subtraction of the lower lip from the upper lip displacement signal. This analysis was carried out to obtain the difference between the upper and lower lip IRED markers as a function of time. Second, the lip aperture trajectories for each nonword were amplitude normalized, which involved subtracting the mean and dividing by the standard deviation. Third, using interpolation, the trajectories of the changes in lip aperture were time-normalized to a fixed record length of 1000 points. Fourth, standard deviation values were calculated at 50 point intervals (2% intervals in relative time) of the normalized waveform. The cumulative sum of the 50 *standard deviations* was computed to obtain a cumulative spatial and temporal coordinative index called lip aperture variability index (Smith & Zelaznik, 2004). The variability in lip aperture is a measure of the trial-to-trial spatial and temporal variability associated with the lip aperture trajectory. Higher lip aperture variability suggests a greater degree of trial to trial variability in inter-articulatory coordination for nonword production, and vice versa. In addition to lip aperture variability, the movement duration in real time was computed as the total duration of the lip aperture signal for each target production and was obtained by measuring the duration of each movement trajectory

of a nonword from start to end point. Figure 2 illustrates an example of the movement trajectories and corresponding lip aperture variability scores for the nonword /*mæbbe b b bibo*/ in a female subject for the delay, gating, and control conditions. Notice the higher score for the delay condition reflecting higher movement variability among the trajectories compared to the gating and control conditions.

Reliability Rating

Inter- and intra-judge reliability for identifying the error trials from the online and offline coding of the responses was derived by dividing the number of agreements by the number of agreements plus disagreements (Sander, 1961). In disfluency counts using this measure, point-by-point reliability values above .80 are considered good, whereas reliability of .90 is considered high (Cordes & Ingham, 1994). This measure yielded an inter-judge reliability rating of 0.98 and an intra-judge reliability rating of 0.94. A range of disagreements between a minimum of zero and a maximum of five trials were identified across the 20 participants.

Statistical Analysis

Several preliminary analyses were conducted on percent correct responses, movement variability, and movement duration to verify normality of distribution and independence of observations. The percent correct responses were arc-sin transformed for further analysis. Both the behavioral and kinematic variables were tested for normality of distribution using the *Kolmogorov – Smirnov* test. Nonsignificant *p* values from this test (percent correct responses, *p* = .18; movement variability, *p* = .20, movement duration, *p* = .20) indicated that the data followed the normal distribution curve. *Levene's test* of homogeneity of variances for all the three dependent measures confirmed the appropriateness of using parametric techniques for analyzing the data (all *p* values were greater than .05). A *Steiger's Z* test of correlation was performed on the three dependent variables to determine if these were intercorrelated. Lack of correlation would allow independent analysis and interpretation of significant effects. This analysis indicated there were significantly different patterns of correlations between the three dependent variables with a significant negative correlation observed only between percent correct responses and movement variability (*Pearson r* = -0.67, *p* = .05). Based on the preliminary analyses, three repeated measures ANOVA were run on the data set. The first and second hypotheses were that speech production errors under delayed auditory feedback and gating would differ from the errors observed in a control, synchronous feedback condition. Both errors and disfluencies were classified as speech errors for the purposes of this analysis. For this analysis, sex of participant was the between-subjects factor while nonword length (six-, 11-syllable) and category (delayed auditory feedback, gating, control) were the within-subjects factors. Percent correct production was the dependent variable. Similar analyses were run to investigate the effects of delayed auditory feedback and gating on speech kinematics. Lip aperture variability and movement duration were the dependent variables.

RESULTS

Percent Correct Responses

Table 1 provides the *Mean* and *SD* values of the percent correct responses for the delay, gating, and control conditions in men and women. A significant main effect of category supported the first hypothesis, $F(2, 36) = 6.45$, $p = 0.004$, $\eta^2 = 0.26$. Post-hoc comparisons (*Fischer's LSD*) revealed a significantly lower percent of correct productions under delayed auditory feedback than the control condition ($p = 0.02$). A similar difference was observed between delayed feedback and gating ($p = 0.02$) thereby failing to confirm the second hypothesis. A non-significant main effect of sex of participant, $F(1, 18) = 1.7$, $p = 0.20$, $\eta^2 = 0.08$, and category x sex interaction, $F(2, 36) = 1.9$, $p = 0.15$, $\eta^2 =$

squared = 0.09, was observed. These results failed to support the third hypothesis of a sex-based difference in performance. Although a significant effect of task complexity was evident as a lower percent of correct productions for the 11-syllable nonwords compared to the six-syllable nonwords, $F(1, 18) = 8.17, p = 0.01, \eta^2 = 0.31$, a nonsignificant category $\times$ length effect, $F(2, 36) = .90, p = 0.41, \eta^2 = 0.04$, failed to support the fourth hypothesis based on task complexity. All other interaction effects were non-significant; length $\times$ sex of participant, $F(1, 18) = 1.7, p = 0.2, \eta^2 = 0.08$; sex $\times$ length $\times$ category, $F(2, 26) = 0.01, p = 0.98, \eta^2 = 0.0006$.

Movement Variability

Table 2 provides the *Mean* and *SD* values of lip aperture variability for the delay, gating, and control conditions in men and women. Once again, a significant main effect of category supported the first hypothesis while failing to confirm the second hypothesis, (*Huynh – Feldt*) $F(2, 34) = 14.07, p = 0.0001, \eta^2 = 0.45$. Post-hoc comparison (*Fischer's LSD*) revealed a significantly higher movement variability for the delay condition compared to the gating ($p = 0.00001$), and the control conditions ($p = 0.0001$). The gating and the control conditions were comparable in movement variability ($p = 0.18$). A nonsignificant main effect of sex of participant, $F(1, 17) = 0.85, p = 0.36, \eta^2 = 0.04$, and a non-significant category $\times$ sex interaction, $F(2, 34) = 2.43, p = 0.103, \eta^2 = 0.12$, failed to support the third hypothesis of sex-based differences in the effects of delayed feedback. However, a significant two-way interaction of sex of participant $\times$ nonword length was observed, $F(1, 17) = 4.5, p = 0.048, \eta^2 = 0.21$ (see Figure 3). Post-hoc comparisons (*Fischer's LSD*) revealed that men and women were comparable in movement variability for the six-syllable nonwords ($p = 0.62$), while women exhibited higher movement variability for the 11-syllable nonwords compared to men ($p = .02$). A significant main effect of length was observed, $F(1, 17) = 153.5, p = 0.0001, \eta^2 = 0.90$. Speakers exhibited a higher movement variability for the longer, 11-syllable nonwords compared to the six-syllable nonwords. However, the results from this analysis failed to support the fourth hypothesis of a task complexity effect, category $\times$ length, $F(2, 34) = 0.2, p = 0.84, \eta^2 = 0.01$. The length $\times$ task $\times$ sex interaction was also non-significant, $F(2, 34) = 1.7, p = 0.33, \eta^2 = 0.02$.

Movement Duration

Table 3 provides the *Mean* and *SD* values of movement duration of lip aperture trajectories for the delay, gating, and control conditions in men and women. A significant main effect of category supported the first hypothesis while failing to support the second hypothesis, $F(2, 34) = 16.3, p = 0.0001, \eta^2 = .49$. Post-hoc comparisons (*Fischer's LSD*) revealed that the movement duration under delayed feedback was significantly slower than under gating ($p = 0.00002$), and control, synchronous feedback conditions ($p = 0.00001$). Movement durations for the gating and the control conditions were comparable ($p > 0.05$). A significant main effect of sex of participant showed that men exhibited slower movement duration compared to women, $F(1, 17) = 8.4, p = 0.01, \eta^2 = 0.33$. However, a non-significant category $\times$ sex interaction, $F(2, 34) = 0.13, p = 0.87, \eta^2 = 0.01$, failed to support the third hypothesis. The analysis revealed a significant main effect of length with participants exhibiting longer movement durations for the 11-syllable nonwords compared to the six-syllable nonwords, $F(1, 17) = 235.7, p = 0.0001, \eta^2 = 0.93$. Furthermore, a significant two-way interaction of length $\times$ category interaction supported the fourth hypothesis, $F(2, 34) = 18.0, p = 0.0001, \eta^2 = 0.51$ (see Figure 5). Post-hoc comparisons (*Fischer's LSD*) revealed all of the differences to be significant except for the movement durations under gating and control conditions for both the six- ($p = 0.69$) and 11-syllable ($p = 0.89$) nonwords, which were comparable.

A significant two-way interaction of length x sex of participant was obtained, $F(1, 17) = 7.7$, $p = 0.013$, $\eta^2 = 0.31$ (see Figure 4). Post-hoc comparisons (*Fischer's LSD*) indicated that while men and women were comparable in movement duration for the six-syllable nonwords ($p = 0.07$), the men were significantly slower in movement duration for the 11-syllable nonwords compared to the women ($p = .001$). Finally, a non-significant sex x category x length interaction was observed in this analysis, $F(2, 34) = 0.03$, $p = 0.97$, $\eta^2 = 0.001$.

DISCUSSION

Effects of Delayed Auditory feedback on Speech Production

Delayed auditory feedback can result in an increase in speech errors and breakdown in speech fluency (e.g., Jones & Striemer, 2007; Corey & Cuddapah, 2008). The findings from the present study support such earlier reports of increase in errors and disfluencies under delayed auditory feedback. Corey and Cuddapah (2008) argued that the delayed signal presents specific challenges to the auditory attention control system, in that the speaker's voice and the delayed feedback signal are identical in all aspects other than timing, and may therefore both be treated as signal (or noise). Delayed auditory feedback can affect speech fluency because the speech production system responds to the delayed signal as it would normally respond to the natural speech signal, thereby creating sequence-inappropriate actions resulting in errors and dysfluencies. The current finding that delayed auditory feedback resulted in more errors supports this interpretation, indicating that participants monitor the feedback for intended content in an ongoing manner and disruption to this feedback can result in errors.

While several studies have investigated momentary compensatory responses of the speech motor system to auditory perturbation (e.g., Mochida, *et al.*, 2010; Cai, *et al.*, 2011), the present study investigated the overall effect of delayed feedback on speech movement coordination. In addition to the momentary compensations reported in earlier studies, delayed auditory feedback has an overall effect of increasing movement variability. This increase could be an epiphenomenon of heightened flexibility of the speech motor system to accommodate the momentary auditory perturbations to achieve the spatio-temporal goals for speech production. This would suggest that the speech motor system uses relevant information from the auditory feedback to fine tune speech movement to achieve end goals. This interpretation would also be in keeping with the earlier studies that have reported of momentary compensatory changes to speech movements in order to accommodate auditory perturbations (Mochida, *et al.*, 2010; Cai, *et al.*, 2011). This is the first study to have demonstrated such a direct effect of delayed auditory feedback on overall movement coordination in continuous speech. This finding is of importance as it supports theories, such as the EXPLAN (Howell, 2004), that attribute a direct effect of delayed auditory feedback on movement kinematics and allows distinction between theoretical accounts (e.g., Howell, 2004 *vs.* MacKay, 1987) that make differing claims about the role of auditory feedback on movement control. The findings also indicate that the measure of spatio-temporal variability, lip aperture variability index, used in the present study is a sensitive measure of the effects of delayed feedback on speech movement kinematics.

Based on certain theoretical accounts (e.g., Disruptive Rhythm Hypothesis; Howell, *et al.*, 1983, the EXPLAN; Howell, 2004) and earlier reports of comparable effects of gating and delayed feedback (e.g., Howell, 2007), it was predicted that gated feedback and delayed feedback would have similar effects on behavioral and kinematic responses including, increase in speech errors, decrease in speech motor coordination and increase in movement duration. To the contrary, the findings from the present study suggested that the two mechanisms, i.e., delayed auditory feedback and gating, work in different ways, at least in

fluent speakers, and do not have similar effects on speech errors, movement variability, and movement duration. In fact, the effect of gating on movement variability was comparable to that of the control condition where no auditory perturbations were applied to speech. This could imply that gating either does not disrupt movement rhythm as does delayed auditory feedback or the severity of the disruption may not be sufficient to have an effect on speech movement kinematics.

The effects of delayed auditory feedback on temporal parameters of connected speech have long been known (Zimmermann, *et al.*, 1988). In the present study, movement duration of the lip trajectories were slower under delayed auditory feedback. This finding corroborates other similar reports of increase in movement duration of speech articulators under delayed feedback (e.g., Sussman & Smith, 1971). Past studies have indicated a reduction in rate of speech with delayed feedback (e.g., Correy & Cuddapah, 2008). The present finding shows that such a reduction in rate is associated with slower speech movement durations and suggests an auditory feedback-based control of movement timing. This seems to be the case despite the long latency of the auditory feedback loop. Slower movement speed under delayed auditory feedback can be interpreted as a strategy for minimizing the auditory mismatch between ongoing utterance and heard utterance. This effect on movement duration was once again specific to delayed feedback and not seen in gating. Furthermore, an effect of task complexity indicated that larger differences were seen in movement duration between the six- and 11-syllable nonwords under delayed auditory feedback compared to the gating and control conditions. Similar effects of task complexity under delayed auditory feedback in speech have been reported, although the stimuli used in such studies were meaningful and varied in cognitive – linguistic complexity (e.g., reading *vs.* spontaneous speech; MacKay, 1987; Cory & Cuddapah, 2008). The present findings suggest that task complexity effects under delayed feedback are not just restricted to meaningful speech, but are also evident in nonwords of varying length that differ primarily in motor complexity. The finding also highlights the importance of task complexity as a determinant of the extent of effects under delayed auditory feedback.

Sex effects

Contrary to earlier reports (e.g., Fukawa, *et al.*, 1988; Cory & Cuddapah, 2008), there were no sex differences in responses to delayed auditory feedback in the present study. Sex differences for movement variability were observed only as a significant group x nonword length interaction, indicating that women compared to men showed higher movement variability for the 11-syllable nonwords. Such differences were not evident for the six-syllable nonwords. A reverse effect was observed in movement duration with the male speakers showing significantly slower movement speed for the 11-syllable compared to the six-syllable nonwords. Overall, the findings from the present study failed to reveal sex-based differences in the processing of delayed feedback.

Conclusions and Limitations

Although earlier studies have investigated momentary compensation in speech movements, to date the overall effects of delayed auditory feedback on speech coordination have not been investigated. Present findings indicate a direct effect of delayed auditory feedback on both behavioral and kinematic performances in a speech task. The findings are of importance as they indicate that delayed auditory feedback induces changes in movement variability and duration in fluent speakers. This reveals that auditory feedback is monitored constantly and is used to update the spatio-temporal characteristics of speech movement trajectories. The findings also support theories of delayed auditory feedback that attribute a role for movement kinematics (e.g., the EXPLAN). Contrary to expectations, gated feedback did not have the same effect on speech movement kinematics. This suggested that the effects

of delayed auditory feedback on movement kinematics are unique and are not mimicked in other types of auditory feedback manipulations.

The present findings need to be interpreted with caution. First, the feedback and gating conditions in the present study may be described as being ‘unnatural’ and therefore the findings can be described as artificial compensations to such perturbations that may not reflect strategies used in everyday situations (for e.g., see Cai, *et al.*, 2011). Further studies are required to understand the effects of delayed auditory feedback on speech movement kinematics in fluent speakers using tasks that induce natural perturbations to speech. Second, the same target nonwords were used in both the practice session and the experiment proper. The familiarity of using the same nonwords and the relatively low skill level required to produce the nonwords could have reducing the magnitude of the effects of delayed auditory feedback and/or gating. Finally, a delay of 200 ms was induced in the feedback signal while a gating frequency of 2 Hz (500 ms “on-off” cycle) was used. It could be argued that the two manipulations are not comparable as one is at syllable rate (i.e., 200 ms) while the other is not. Future studies need to compare different values of delay and gating to find if the observed effects under these altered feedback conditions are comparable at any specific rates.

Acknowledgments

This study was funded by an R03 DC 010047 from NIDCD. The author wishes to acknowledge Katie Lauritzen, Christine McConnell, and Carol June Leonard for data collection and analysis, Dr. Edward Carney for technical assistance.

References

- Antipova E, Purdy S, Blakeley M, WILLIAMS S. Effects of altered auditory feedback (AAF) on stuttering frequency during monologue speech production. *Journal of Fluency Disorders*. 2008; 33:274–290. [PubMed: 19328980]
- Borden GJ. Interpretation of research on feedback interruption in speech. *Brain and Language*. 1979; 7:307–319. [PubMed: 455050]
- Cai S, Ghosh SS, Guenther FH, Perkell JS. Focal manipulations of formant trajectories reveal a role of auditory feedback in the online control of both within-syllable and between-syllable speech timing. *The Journal of Neuroscience*. 2011; 31(45):16483–16490. [PubMed: 22072698]
- Chase RA. Effect of delayed auditory feedback on the repetition of speech sounds. *Journal of speech and hearing Disorders*. 1958; 23:583–590. [PubMed: 13588683]
- Cordes AK, Ingham RJ. The reliability of observational data: II. Issues in the identification and measurement of stuttering events. *Journal of Speech and Hearing Research*. 1994; 37:279–294. [PubMed: 8028309]
- Corey DM, Cuddapah VA. Delayed auditory feedback effects during reading and conversation tasks: Gender differences in fluent adults. *Journal of Fluency Disorders*. 2008; 33:291–305. [PubMed: 19328981]
- Fairbanks G. Selective vocal effects of delayed auditory feedback. *Journal of Speech and Hearing Disorders*. 1955; 20:333–348. [PubMed: 13272227]
- Fukawa T, Yoshioka H, OZAWA E, Yoshida S. Difference of susceptibility to delayed auditory feedback between stutterers and nonstutterers. *Journal of Speech and Hearing Research*. 1988; 31:475–479. [PubMed: 3172765]
- Goldiamond, I. Stuttering and fluency as manipulatable operant response classes. In: Krasner, L.; Ullman, L., editors. *Research in behaviour modification*. New York, NY: Holt, Rinehart and Winston; 1965. p. 106-156.
- Harper L, Houle K. Speech self-monitoring in preschoolers: The effects of delayed auditory feedback on recitation. *Perceptual and Motor Skills*. 2000; 90:1157–1170. [PubMed: 10939063]

- Howell P. Assessment of some contemporary theories of stuttering that apply to spontaneous speech. *Contemporary Issues in Communicative Sciences and Disorders*. 2004; 39:122–139.
- Howell P. The effects of gated speech on the fluency of speakers who stutter. *Folia Phoniatica et Logopaedica*. 2007; 59:250–255. [PubMed: 17726328]
- Howell P, Archer A. Susceptibility to the effects of delayed auditory feedback. *Perceptual Psychophysics*. 1984; 36:296–302.
- Howell P, Powell DJ, Khan I. Amplitude contour of the delayed signal and interference in delayed auditory feedback tasks. *Journal of Experimental Psychology: Human Perception & Performance*. 1983; 9:772–784.
- Howell P, Sackin S. Timing interference in altered listening conditions. *Journal of the Acoustical Society of America*. 2002; 111:2842–2852. [PubMed: 12083218]
- Jones JA, Striemer D. Speech disruption during delayed auditory feedback with simultaneous visual feedback. *Journal of the Acoustic Society of America*. 2007; 122(4):135–141.
- Kaspar K, Rubeling H. Rhythmic Versus Phonemic Interference in Delayed Auditory Feedback. *Journal of Speech, Language, and Hearing Research*. 2011; 54:932–943.
- Lane H, Tranel B. Lombard sign and role of hearing in speech. *Journal of Speech and Hearing Research*. 1971; 14:677–709.
- Lee, Bs. Some effects of side-tone delay. *Journal of the Acoustical Society of America*. 1950; 22:639–640.
- Lincoln M, Packman A, Onslow M. Altered auditory feedback and the treatment of stuttering: A review. *Journal of Fluency Disorders*. 2006; 31:71–89. [PubMed: 16750562]
- Lincoln M, Packman A, Onslow M, JONES M. An experimental investigation of the effect of altered auditory feedback on the conversational speech of adults who stutter. *Journal of Speech Language and Hearing Research*. 2010; 3(5):1122–31.
- Lotzmann G. Zur Anwendung variierter Verzögerungszeiten bei Balbuties. *Folia Phoniatica*. 1961; 13:276–310. [PubMed: 14466801]
- Mackay, DG. *The organization of perception and action*. New York: Springer-Verlag; 1987.
- Miller GA, Licklider JCR. The intelligibility of interrupted speech. *Journal of the Acoustic Society of America*. 1950; 22:167–173.
- Mochida T, Gomi H, Kashino M. Rapid change in articulatory lip movement induced by preceding auditory feedback during production of bilabial plosives. *PLoS ONE*. 2010; 5(11):e13866. [PubMed: 21079783]
- Moeleants, D.; Demey, M.; Leman, M. Performing music with delayed auditory feedback. In: Louhivouri, J.; Eerola, T.; Saarikallio, S.; Himberg, T.; Eerola, P-S., editors. *Proceedings of the 7th Triennial Conference of European Society for the Cognitive Sciences of Music (ESCOM 2009)*. Finland: Jyväskylä; 2009. p. 348-352.
- Perkins WH, Bell J, Johnson L, Stochs J. Phone rate and the effective planning time hypothesis of stuttering. *Journal of Speech and Hearing Research*. 1979; 22:747–755. [PubMed: 513683]
- Pfordresher PQ. Auditory feedback in music performance: Evidence for a dissociation of sequencing and timing. *Journal of Experimental Psychology: Human Perception and Performance*. 2003; 29:949–964. [PubMed: 14585016]
- Pfordresher PQ, Benitez B. Temporal coordination between actions and sound during sequence production. *Human Movement Science*. 2007; 26:742–756. [PubMed: 17764772]
- Pfordresher PQ, Dalla Bella SD. Delayed Auditory Feedback and Movement. *Journal of Experimental Psychology*. 2011; 37(2):566–579. [PubMed: 21463087]
- Purcell DW, Munhall KG. Compensation following real-time manipulation of formants in isolated vowels. *Journal of Acoustic Society of America*. 2006; 119:2288–2297.
- Siegel GM, Fehst CA, Garber SR, PICK HL Jr. Delayed auditory feedback with children. *Journal of Speech Hearing Research*. 1980; 23(4):802–813. [PubMed: 7442213]
- Smith A, Johnson M, MCGILLEM C, Goffman L. On the assessment of stability and patterning of speech movements. *Journal of Speech Language and Hearing Research*. 2000; 43(1):277–286.
- Sussman HM, Smith KU. Jaw movements under delayed auditory feedback. *Journal of the Acoustical Society of America*. 1971; 50:685–691. [PubMed: 5096504]

- Swink S, Stewart A. The effect of gender on the N1–P2 auditory complex while listening and speaking with altered auditory feedback. *Brain & Language*. 2012; 122(1):25–33. [PubMed: 22564750]
- Walsh B, Smith A, Weber-Fox C. Short-term plasticity in children’s speech motor systems. *Developmental Psychobiology*. 2006; 48(8):660–674. [PubMed: 17111401]
- Zimmerman G, Brown C, Kelso JAS, Hurtig R, Forrest K. The association between acoustic and articulatory events in a delayed auditory feedback paradigm. *Journal of Phonetics*. 1988; 16:437–451.

Figure 1.
The instrumental set-up for the delayed feedback and gating protocol experiment.

Figure 2.

An example of movement trajectories and lip aperture variability scores for the nonword /*mæbbe b b bibo*/ for the delay, gating, and control conditions in a female subject.

Figure 3.
Lip aperture variability (and SE) for men and women across two nonword lengths.

Figure 4. Movement duration (in sec) and SE for men and women across two nonword lengths.

Figure 5. Movement duration (in sec) and SE across two nonword lengths for delayed feedback, gating, and control conditions.

Table 1

Mean and SD of percent correct productions for men and women across DAF, Gating, and Control conditions.

Group	length	task	Mean	SD
Females	6- syllable	Control	87.4	9.05
		Gating	88.9	8.8
		DAF	84.1	18.0
	11-syllable	Control	86.8	14.2
		Gating	88.7	13.7
		DAF	80.4	17.4
Males	6-syllable	Control	94.1	4.3
		Gating	93.2	7.6
		DAF	94.2	6.2
	11-syllable	Control	90.2	7.9
		Gating	90.2	5.4
		DAF	88.1	9.4

Table 2

Mean lip aperture variability and SD for the two groups across DAF, Gating, and Control conditions.

Group	length	task	Mean	SD
Females	6-syllable	Control	20.956	6.3
		Gating	19.134	4.5
		DAF	24.986	4.3
	11-syllable	Control	27.221	5.2
		Gating	24.444	4.2
		DAF	30.245	4.5
Males	6-syllable	Control	19.994	7.3
		Gating	20.110	4.3
		DAF	22.040	3.9
	11-syllable	Control	23.790	6.7
		Gating	23.714	5.1
		DAF	26.528	3.2

Table 3

Mean movement duration and SD for the two groups across DAF, Gating, and Control conditions.

Group	length	task	Mean	SD
Females	6-syllable	Control	1.9	0.4
		Gating	1.9	0.3
		DAF	2.4	0.9
	11-syllable	Control	3.9	0.7
		Gating	3.8	1.9
		DAF	5.0	1.1
Males	6-syllable	Control	2.8	0.7
		Gating	2.9	0.9
		DAF	3.4	1.1
	11-syllable	Control	5.8	1.4
		Gating	5.9	1.7
		DAF	7.0	2.2