2024-07-25

Exploring the Effects of Delayed Auditory Feedback on Speech Kinematics: A Comparative Analysis of Monologue Speech and Tongue Twisters

Abbey Corinne Persons
Brigham Young University

Exploring the Effects of Delayed Auditory Feedback on Speech Kinematics:

A Comparative Analysis of Monologue Speech and Tongue Twisters

Abbey Corinne Persons

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Master of Science

Christopher Dromey, Chair
Connie Summers
Dallin Bailey

Department of Communication Disorders

Brigham Young University

ABSTRACT

Exploring the Effects of Delayed Auditory Feedback on Speech Kinematics:
A Comparative Analysis of Monologue Speech and Tongue Twisters

Abbey Corinne Persons
Department of Communication Disorders, BYU
Master of Science

This study investigated the effects of delayed auditory feedback (DAF) on speech kinematics during tongue twisters and monologues. Participants were 20 adults (10 men, 10 women) aged 18–29 with typical speech and hearing abilities. A smartphone app provided DAF latencies of 50 ms, 100 ms, and 150 ms. Kinematic measures were made of peak speed, stroke distance, and hull area for the tongue front, jaw, and lower lip under typical and the three DAF conditions. Results indicated that DAF significantly reduced peak speed and stroke distance for all articulators during tongue twisters ($p < .01$), with the effect magnitude increasing with longer delays. No significant DAF effects were observed in monologues ($p > .05$). Sex differences were noted, with women showing higher speeds and longer stroke distances across both tasks ($p < .05$). These findings suggest that DAF disrupts motor performance, particularly in structured tasks, and that sex differences are present in speech kinematics. Future research could explore the perceptual impacts of DAF and the cognitive load associated with speech under altered feedback conditions.

Keywords: speech kinematics, articulation, delayed auditory feedback

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to my committee members, Dr. Dromey, as chair, with Dr. Summers and Dr. Bailey, for their invaluable guidance, support, and encouragement throughout this research journey. Their expertise and insights were instrumental in shaping this thesis.

I am profoundly grateful to my family and friends for their unwavering support and patience during this process. Special thanks to my mom, dad, and sisters for their constant encouragement, understanding, and assistance.

Lastly, I want to extend my appreciation to all of the participants who took part in this study. Their cooperation and willingness to contribute their time and effort are greatly appreciated.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

DESCRIPTION OF THESIS STRUCTURE AND CONTENT

This thesis, *Exploring the Effects of Delayed Auditory Feedback on Speech Kinematics: A Comparative Analysis of Monologue Speech and Tongue Twisters*, is presented in a hybrid format designed to meet the length and style requirements for submission to research journals in communication disorders. The initial sections adhere to the university's submission guidelines, and the annotated bibliography can be found in Appendix A. The approved consent form can be found in Appendix B. The Institutional Review Board approval letter is in Appendix C.

Introduction

Auditory feedback plays a crucial role in the production and control of speech because it allows speakers to monitor and adjust their performance. Through this process, speakers can detect errors such as mispronunciations or other deviations from their intended speech patterns. The feedback enables self-correction and the maintenance of precision over time, ensuring that speech remains clear and accurate. Auditory feedback is critical for the development of speech sound production. It helps individuals learn and refine the articulatory movements and muscle coordination required for accurate speech sound production. This is particularly important during language acquisition in children. It can also influence prosody and intonation of speech, which convey meaning, emotion, and emphasis (Katseff et al., 2012).

The use of cochlear implants can shed light on the importance of auditory feedback in speech production. Cochlear implants have a profound impact on speech production, particularly for individuals with severe to profound hearing loss. These sophisticated devices work by bypassing damaged or non-functioning parts of the inner ear and directly stimulating the auditory nerve, allowing recipients to perceive sound. As a result, individuals who receive cochlear implants often experience significant improvements in their ability to understand and produce speech. The enhanced auditory input enables them to access a broader range of sounds, aiding in the development of clearer and more accurate speech. Cochlear implants facilitate the recognition of subtle nuances in pitch, tone, and rhythm, which are crucial components of effective communication. Gains in speech production can be found as early as 1 month after tune-up of the device (Te et al., 1996).

One study by Svirsky et al. (1992) examined the speech patterns of four individuals before and after receiving cochlear implants. It found abnormalities in pre-implant speech, with variations among individuals. Two women, deafened in adulthood, showed trends toward normalized speech post-implant, while two men, deafened in childhood, had more complicated outcomes. Changes in speech parameters like airflow, fundamental frequency, and vowel duration were observed post-implant. The study suggests that perceptual gains and prior linguistic experience influence speech production changes (Svirsky et al., 1992).

The Svirsky et al. (1992) study is significant for understanding the importance of auditory feedback because it identified abnormalities in the speech patterns of post-lingually deafened individuals before receiving cochlear implants. This suggests that auditory feedback plays a crucial role in shaping speech production, as these abnormalities likely stem from deficits in perceiving and processing auditory information. After receiving cochlear implants, these individuals showed changes in various speech parameters, such as airflow, fundamental frequency, and vowel duration. These changes suggest that auditory feedback provided by the implants influenced speech production, leading to adjustments aimed at aligning speech output with perceived auditory targets.

Altered auditory feedback (AAF) refers to the manipulation of an individual's perception of their own speech sounds in real-time during production. This experimental technique has been widely used to investigate the relationship between auditory perception and speech production. When individuals hear their own speech with modified pitch, delay, or other alterations, it can have a significant impact on their ability to produce fluent and intelligible speech. Experimental evidence suggests that AAF can reduce stuttering during oral reading in controlled laboratory conditions. However, the effectiveness of AAF in conversational speech remains uncertain (Lincoln et al., 2006).

Studies on AAF have demonstrated its potential to induce changes in various aspects of speech, providing insights into the mechanisms underlying speech motor control and feedback processes. For example, alterations in pitch perception can lead to compensatory adjustments in vocal pitch to maintain a sense of normalcy in one's own auditory feedback. These experiments contribute valuable information to our understanding of how the brain integrates auditory information with motor control during speech production (Hain et al., 2000).

Delayed auditory feedback (DAF) is a type of altered auditory feedback that is used in research and speech therapy where an individual's speech is played back to them with a delay. This delayed auditory feedback can cause disruptions in an individual's normal speech patterns. When individuals hear their own voice with a short delay, it can lead to disfluencies, such as repetitions of sounds or words, prolonged speech sounds, and altered speaking rates (Jones & Striemer, 2007).

While DAF can make typical speakers less fluent, it can be used as a therapy technique for individuals who stutter. One study found that combining DAF with prolonged speech reduced a speaker's rate and increased his intelligibility scores (Dagenais et al., 1998). DAF can improve fluency by providing feedback that disrupts typical stuttering patterns and encourages smoother speech production (Chon et al., 2021). One of the leading theories for improved fluency with DAF in stutterers is that introducing a slight delay in auditory feedback disrupts the normal feedback loop between speaking and hearing. This disruption can influence speech motor control and coordination, which can result in fewer disfluencies.

The study of DAF holds significance in the field of speech research, as it constitutes a means to understanding and addressing speech disorders. DAF offers a unique window into the intricate dynamics of speech production and fluency.

Changes in speech can be perceived by listeners if they are significant enough. However, there are small changes in our speech movements that can only be detected using instrumentation. Speech movements can be effectively measured using sensors placed strategically in and around the mouth. These sensors, often part of advanced speech analysis systems, play a crucial role in capturing the intricate movements involved in the process of speech production. One type of system employed for this purpose is the electromagnetic articulograph, which uses small sensors attached to the tongue, lips, and jaw. These sensors detect and record the subtle changes in position and orientation of these articulators during speech, providing valuable data for analyzing speech patterns and articulatory dynamics.

To more fully understand the effect of DAF on speech, it is helpful to explore it in multiple contexts. Reading a script and engaging in a monologue present distinct challenges and variations in terms of speech movements. When individuals read from a script, the speech movements are often more controlled and deliberate. The speaker has the opportunity to review and rehearse the text, leading to a more planned and structured delivery. In scripted scenarios, the movements of articulators such as the tongue, lips, and jaw could potentially be more consistent, as the speaker talks without the spontaneity and uncertainty inherent in natural conversation or monologue.

Conversely, in a spontaneous monologue, the speaker must respond in real-time to the flow of their own speech, necessitating quick adjustments in pronunciation, pitch, and pacing. Unlike reading from a script, conversational speech movements are influenced by factors like turn-taking, the emotional tone of the interaction, and the need for on-the-fly decision-making. These factors contribute to a more fluid and adaptive use of articulators, reflecting the dynamic nature of interactive communication.

Choosing to include spontaneous speech in a study can have pros and cons as it relates to balancing experimental control and ecological validity. These are two important concepts in an experimental design and represent different considerations that must be balanced. Experimental control refers to the extent to which a researcher can manipulate and control the variables in an experiment to ensure that changes in the dependent variable are due to the manipulation of the independent variable and not to other extraneous factors. Experimental control is crucial for establishing cause-and-effect relationships. By carefully controlling and manipulating variables, researchers can isolate the impact of the independent variable on the dependent variable, minimizing the influence of confounding variables. A highly controlled study may use contrived stimuli that are repeated multiple times to allow the same measures to be compared under different conditions.

Ecological validity, on the other hand, refers to the extent to which the findings of a study generalize or apply to real-world, everyday situations. In other words, the experimental conditions and procedures reflect the complexity and dynamics of the real world. Ecological validity is essential for the external validity or generalizability of research findings. If a study lacks ecological validity, its applicability to real-life situations may be limited, and the findings may not accurately represent how people would behave or respond outside the laboratory setting. An example of high ecological validity would include observing children's social interactions on a playground or in a classroom. The behaviors observed in these settings are likely to be more representative of typical social interactions than those in a controlled lab environment.

On the other hand, an example of low ecological validity would be an experimental task where participants learn a list of words in isolation within a controlled environment. In real life, vocabulary acquisition is often context-dependent, influenced by various cues and environmental factors.

While reading a script can sometimes produce more consistent speech, certain scripts can prove to be more challenging than others. Tongue-twisters, for example, are sequences of words or phrases that are designed to be difficult to articulate rapidly and accurately. They often consist of a series of similar sounds or phonemes, requiring the speaker to pronounce them quickly without stumbling over the words. Classic examples include phrases like "She sells seashells by the seashore" or "Peter Piper picked a peck of pickled peppers." The primary purpose of tongue twisters is to challenge the speaker's ability to articulate sounds, syllables, and words.

In order to measure differences between speech with DAF and without DAF as well as the effects of coarticulation, several techniques have been used to quantify speech movements. Traditional kinematic measures used in speech analysis include displacement, peak velocity, and acceleration during articulatory gestures. In speech kinematics, displacement refers to the distance an articulator travels from one position to another. For example, the displacement of the tongue during the production of a syllable involves its movement from one phoneme to the next. Velocity in speech kinematics represents the rate of change in displacement of articulators over time. It provides information about how quickly articulators move during speech production. For instance, the velocity of lip movement can be analyzed to understand the rapidity of articulation in forming different speech sounds. Acceleration is the rate of change in velocity. In speech kinematics, it describes how quickly the velocity of articulator movement changes during speech production. Analyzing acceleration can offer insights into the dynamic aspects of speech articulation.

An alternative to traditional kinematic measures like displacement and velocity is to use stroke measures to define speech movements. A stroke is defined as the movement that occurs between two consecutive local minima in the speed history of an articulator. This method is applied to segment the movements of the articulators during connected speech; it can be applied to any length of recording (Tasko & Westbury, 2002).

Stroke measures can provide global, averaged metrics reflecting the control of articulatory movements and document the effects of interventions. When investigating the effects of interventions, such as speech therapy techniques or technological interventions like DAF, stroke measures can help quantify overall changes in articulatory behavior. This is essential for understanding how interventions influence speech kinematics in natural conversational contexts because traditional kinematic measures reflect the features of specifically selected phonemes, whereas strokes can be applied to any speech recording, no matter how complex or variable the stimuli.

Stroke measures such as average stroke distance, stroke duration, peak stroke speed, and boundary speed can be used to compare different speech samples. This is helpful when comparing the small changes our articulators make when under different conditions such as DAF, no DAF, reading a script, or speaking in a monologue. Stroke distance is the average total path length covered by an articulator during a series of speech gestures. Stroke duration is the average time taken to complete each speech gesture. It measures the temporal aspect of the movement, which may change under different speaking conditions.

Peak stroke speed is the average maximum velocity reached by an articulator over a series of speech movements. It provides information about the fastest point during the gesture and can be indicative of the force or energy applied during speech. Boundary speed or the speed minimum refers to the average lowest speed during a series of speech gestures.

Using kinematics to measure the effects of DAF on speech production may reveal distinct patterns of motor adjustments in response to altered auditory feedback. Specifically, it is hypothesized that individuals will exhibit changes in articulatory kinematics, such as alterations in speech rate and movement variability, when speech is subjected to varying degrees of DAF. Furthermore, it is anticipated that these kinematic changes will be influenced by factors such as the magnitude of the delay, individual differences in speech motor control, and the nature of the speech task. By employing kinematic analysis, this research aims to explain the underlying mechanisms of how DAF affects speech production.

Method

Participants

The study included 20 young adults who were speakers of Standard American English and had no history of speech, language, or hearing disorders, as determined during an initial interview. This group consisted of 10 men aged 18 to 29 and 10 women aged 18 to 25, all of whom were recruited from the university community. Before the experiment, participants received an explanation of the research and completed a consent form approved by the Institutional Review Board at Brigham Young University.

Instrumentation

Multiple sets of data were gathered during each recording session as part of a larger study, although only a portion of these data analyzed for this study. All recordings took place in a single-walled sound booth, with the participant positioned 50 cm from a condenser microphone (AKG C2000B). To facilitate software-based measurement of speech intensity in dB sound pressure level (SPL) based on the microphone signal, a calibration vowel was recorded at a distance of 50 cm from a sound level meter.

Articulatory kinematic data were captured using an NDI Wave electromagnetic articulograph from Northern Digital Inc. This device recorded data from 3 mm sensor coils, including two reference sensors affixed to an eyeglass frame to establish the coordinate system and correct for head movements. To track articulatory movement two coils were positioned at midline on the vermillion borders of the upper (UL) and lower lips (LL), one at the tongue front (TF), one at the middle of the tongue (TM). One coil was attached to the central incisors of the mandible to measure jaw movement (J). The sensor coils were securely attached using cyanoacrylate dental adhesive, with a small piece of Stomahesive (ConvaTec, Inc.) applied to the jaw sensor for the protection of tooth enamel.

Outside the sound booth, a computer recorded the x, y, and z positions of each coil with the Wavefront software application (Northern Digital Inc., 2017). The movement data were sampled at a rate of 100 Hz, while the time-aligned audio signal was sampled at a rate of 22,050 Hz.

Procedure

After coils were attached to the articulators, the experimenter engaged the participant in conversation for five minutes to help them habituate to the sensation of speaking with sensors attached to minimize learning effects during data collection. After this adaptation period, headphones were placed on the participant's ears for them to hear the DAF of their speech.

Participants read aloud the following tongue twister four times under each DAF condition: "If a dog chews shoes, whose shoes does he choose." After four repetitions under one delay condition, the participant was exposed to a new delay which included 0-ms, 50 ms, 100 ms, and 150 ms. Studies have shown that longer auditory delay settings can lead to greater reductions in speech rate for both people who stutter and speakers with Parkinson's disease (Brendel et al., 2004). Additionally, research has revealed that a 150 ms delay setting is particularly effective for reducing speech rate and improving intelligibility (Brendel et al., 2004). Based on these findings, the current study with typical speakers included delay settings of 50 ms, 100 ms, and 150 ms to explore the effects on speech movement patterns, leveraging the insights from previous studies. To counteract any practice effects with the tongue twister task, the DAF conditions were presented in a randomized order. The participant also engaged in a monologue in response to the following two prompts: "Tell me about your favorite vacation you've ever taken," and "Tell me about your favorite class you've ever taken." One monologue was conducted without DAF and the other monologue was conducted with DAF (150 ms delay). The monologue prompts and levels of DAF stimuli were presented in a randomized order. Four DAF conditions were chosen for the tongue twister task to comprehensively explore the impact of varying delays on complex speech patterns, while two DAF conditions were selected for the monologue task to focus on the effects of the longest delay in more natural, spontaneous speech.

The tongue twister and monologue stimuli were chosen to compare the effects of DAF on speakers producing spontaneous language versus reading from a script. Each stimulus was presented on a screen for the participants to read after instruction.

Data Analysis

Segmentation

The initial step involved rough segmentation of the kinematic recording for each participant into distinct tokens corresponding to individual stimuli, accomplished with a custom Matlab application (Mathworks, 2023). Subsequently, a second custom Matlab application was employed for finer segmentation of each token. Cursors were placed at the beginning and end of the tongue twister utterance to segment.

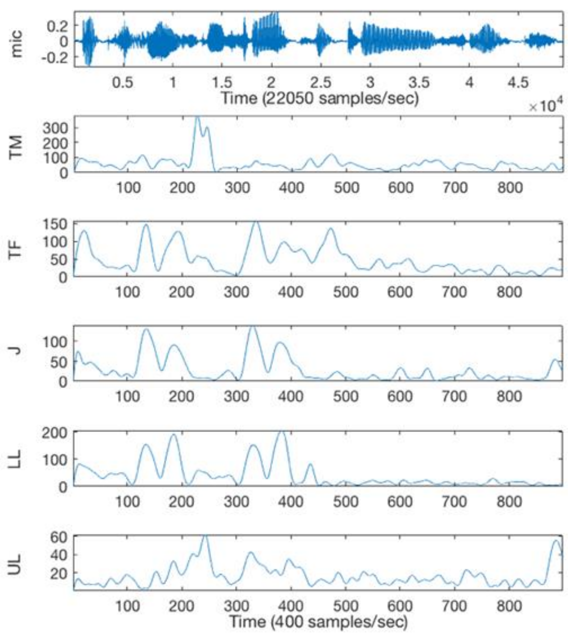
The analysis included measuring a 30-second segment from each monologue with a custom Matlab application. For the tongue twister, several measures were calculated, including sentence duration and metrics reflecting articulator movement strokes, as outlined below. The stroke metrics were also computed for the monologue sample.

Stroke Metrics

Stroke metrics were computed from speed plots in Matlab, which were based on the Euclidean distance or segment length between adjacent X and Y points in the kinematic recording. Thus, the speed record reflected a change in position over time irrespective of direction. A stroke was defined as the movement between one speed minimum and the next throughout the speed history (see Figure 1).

Figure 1

Audio and Articulator Speed Plots (mm/s) for the Stimulus Sentence

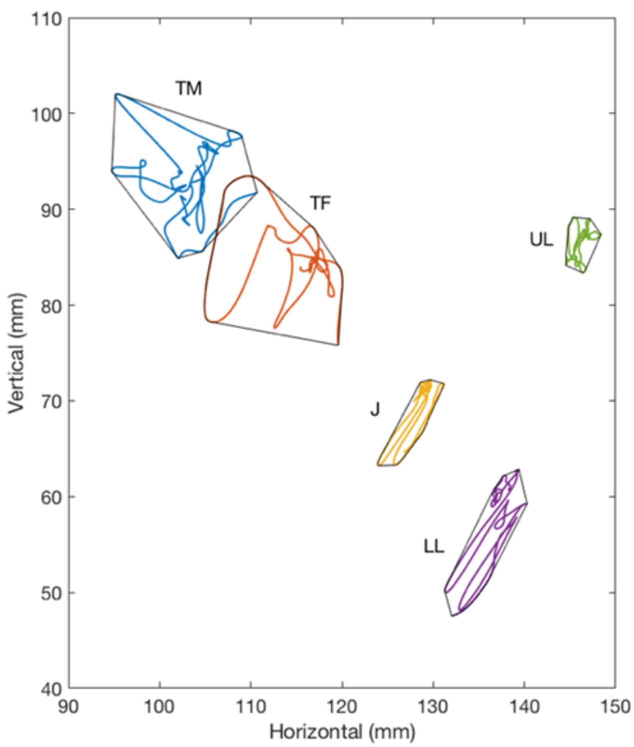


Note. TM= tongue mid, TF = tongue front, J = jaw, LL = lower lip, UL= upper lip.

Stroke measures were averages for the entire analysis segment and included several kinematic data metrics: the stroke count (total number of articulatory strokes), onset speed, peak speed, mean speed, stroke distance, and stroke duration. Additionally, the 2-dimensional area in mm^2 covered by all articulatory movements in the vertical and horizontal directions during the target sentence (computed using the convex hull operation in Matlab) was determined. These kinematic stroke metrics provide insight into articulator movement during continuous speech, presenting a comprehensive view of speech system movements averaged across entire utterances, unlike traditional point measures that focus on single movements for specific syllables or phonemes (see Figure 2).

Figure 2

Position and Convex Hull of all Sensors During the Stimulus Sentence.



Statistical Analysis

A linear mixed model (LMM) analysis in SPSS 29 was conducted to test for differences in the dependent variables across the experimental conditions. DAF latency (0 ms, 50 ms, 100 ms, and 150 ms) was the fixed factor for the tongue twister task. Presence or absence of DAF was the fixed factor for the monologue task. Speaker sex was a fixed factor for both tasks, and speaker was the random factor. Bonferroni-corrected pairwise multiple comparisons were used to test for significant differences between no DAF and the three latency conditions for the tongue twister task.

Results

This study examined the effect of delayed auditory feedback on tongue twister and monologue speech tasks. From the larger dataset collected during the experiment, three articulators were selected for reporting here. They were the tongue-front (TF), jaw (J), and lower lip (LL). The kinematic measures that were selected to report are stroke count, peak speed, stroke distance, and hull area. These articulators and measures were chosen in order to report the primary findings from a large and complex dataset.

Descriptive statistics for the effects of DAF on the speech kinematic measures are reported in Table 1, and for male and female participants in Table 2. Tables 3 and 4 report the results from the linear mixed model ANOVA testing.

Tongue Twister Task

Stroke Count

The stroke count did not change significantly across DAF conditions or differ between men and women for any of the articulators.

Peak Speed

There was a significant main effect of DAF condition on peak speed for the TF, J, and LL (see Table 3). Pairwise comparisons revealed that the peak speed was lower for each DAF latency condition for the TF ($p < .001$). For the J, pairwise comparisons showed a significant decrease in peak speed for the 100 ms delay ($p = .035$) and the 150 ms delay ($p = .002$). For the LL, pairwise comparisons showed a significant decrease in peak speed for the 50 ms delay ($p = .034$), for the 100 ms delay ($p = .027$) and the 150 ms delay ($p = .001$). For the J peak speed, there was a significant difference between male and female speakers (see Table 3), with a higher speed for women (Table 2).

Stroke Distance

A significant main effect of DAF condition on TF stroke distance was found. Pairwise comparisons revealed that TF stroke distance was significantly shorter for each DAF latency condition compared to the no DAF condition. Specifically, TF distance decreased significantly for the 50 ms delay (p = .009), the 100 ms delay (p = .019), and the 150 ms delay (p = .001). For J stroke distance, there was a significant main effect of DAF condition, with pairwise comparisons showing that J distance was significantly shorter for the 150 ms delay compared to the no DAF condition (p = .030). Moreover, a significant effect of sex on J distance was found, indicating that J distance was significantly shorter for men compared to women (p = .014). Similarly, a significant main effect of DAF condition was observed on LL distance, with pairwise comparisons showing that LL distance was significantly shorter for the 150 ms delay compared to the no DAF condition (p = .010). Additionally, a significant effect of sex on LL distance was found, indicating that LL distance was significantly shorter for men compared to women (p = 0.019).

Hull Area

There was no significant effect of DAF on hull area for any of the articulators. However, a significant effect of sex was observed on J hull area, which was significantly smaller for men compared to women (p = .022). Similarly, a significant effect of sex was observed on LL hull area, which was significantly smaller for men compared to women (p = .019).

Monologue Task

Stroke Count

There were no significant effects of DAF on stroke count for any of the articulators. Sex had a significant effect on TF, J, and LL stroke counts. Men had significantly higher stroke counts than women for all three articulators ($p < .001$).

Peak Speed

There were no significant effects of DAF on peak speed for any of the articulators. There was a significant effect of sex on J peak speed, with men exhibiting significantly lower speeds compared to women ($p = .021$). Similarly, for LL peak speed, men had significantly lower speeds than women ($p = .015$).

Stroke Distance

There were no significant effects of DAF on stroke distance for any of the articulators. A significant effect of sex on J stroke distance was observed, which was significantly shorter for men compared to women ($p = .010$). Likewise, for LL stroke distance, men exhibited significantly shorter distances than women ($p = .008$).

Hull Area

There were no significant effects of DAF on hull area for any of the articulators. There was a significant effect of sex on J hull area, with men having significantly smaller areas compared to women ($p = .049$). Similarly, for LL hull area, men exhibited significantly smaller areas than women ($p = .036$).

Discussion

This study investigated the effects of delayed auditory feedback (DAF) on speech tasks involving tongue twisters and monologues, focusing on the tongue-front (TF), jaw (J), and lower lip (LL). For the tongue twister task, DAF significantly reduced peak speed and stroke distance for all three articulators, with these effects varying with DAF latency. Sex differences were noted, with women generally showing higher speeds and longer stroke distances. In the monologue task, DAF had no significant effects on peak speed, stroke distance, or hull area, but sex differences persisted, with men typically having lower stroke speeds, shorter stroke distances, and smaller hull areas.

A study by Nissen et al. (2007) investigated kinematic differences in proficient bilingual speakers as they produced their first language (L1, which was Korean or Spanish) and second language (L2, which was English) and how these differences related to the perceived strength of their accent in L2. By analyzing the peak and average tongue stroke speeds of Korean and Spanish speakers across three speaking tasks (monologue, picture description, reading), the authors found that speakers exhibited slower stroke speeds when speaking in L2. This slower speed correlated with higher accent ratings and lower speech-to-pause ratios, suggesting that the kinematic differences were perceptually significant. The authors suggested that the increased effort required for articulatory flexibility and automaticity in L2 accounted for these differences, highlighting the lower level of automaticity when speaking in a non-native language.

In the Nissen et al. (2007) study, significant differences in peak and average tongue stroke speeds between L1 and L2 were found, with slower speeds in L2 suggesting increased effort and decreased articulatory automaticity. This is comparable with the findings of the current study on the effect of delayed auditory feedback (DAF) on tongue twister and monologue tasks, where peak speeds for the tongue-front (TF), jaw (J), and lower lip (LL) were significantly reduced under DAF conditions, suggesting a similar increase in effort and disruption in motor performance due to external feedback manipulation. However, unlike the bilingual study, which reported slower stroke speeds without changes in stroke distance, the current findings revealed that stroke distances for TF, J, and LL decreased significantly under DAF, particularly at longer delay intervals (100 ms and 150 ms). This indicates that DAF not only affects speed but also the extent of articulatory movement.

Both studies highlighted the influence of different speech tasks on kinematic measures. The Nissen et al. (2007) study noted faster stroke speeds during paragraph reading compared to monologue and picture description tasks, which the authors attributed to the familiarity and structured nature of reading. Similarly, in the present study, while DAF significantly affected peak speeds and stroke distances in tongue twisters (a more structured and challenging task), these effects were not observed in monologues. This aligns with the notion that structured tasks (e.g., reading or tongue twisters) may impose greater articulatory demands, thus being more susceptible to external disruptions like DAF or L2 speaking.

The current study found significant sex differences in peak speeds, stroke distances, and hull areas for J and LL, with women generally showing higher speeds and larger distances. These differences were consistent across both DAF and non-DAF conditions, suggesting inherent sex-based differences in speech kinematics. This finding emphasizes the importance of considering sex as a variable in speech production studies, which may provide further insights into the neuromuscular and articulatory characteristics across different populations.

Anatomical differences between men and women, such as men having larger vocal tracts and articulators, might lead one to expect that men would exhibit larger hull areas during speech production. The hull area, representing the spatial extent of articulatory movements, could reasonably be presumed larger in men due to their generally larger anatomical structures. However, the findings of the current study contradict this assumption, showing that women had larger hull areas for both the jaw and lower lip compared to men. This counterintuitive finding could be attributed to women speaking more precisely, which involves greater and more controlled movements of the articulators. Women's speech precision may necessitate more extensive articulatory excursions, resulting in larger hull areas despite their relatively smaller anatomical structures. This precision in speech could be linked to sociolinguistic factors, where women are often observed to articulate more clearly and carefully. A study by Taylor et al. (2020) reported that women had higher spectral mean and kurtosis values for /s/ than men. This was hypothesized to be due to differences in vocal tract anatomy, such as a more anterior and/or deeper tongue groove in women. While neither Taylor et al. (2020) or the present study involved perceptual ratings of speech precision, the data from both studies point to articulatory differences that appear to reflect greater precision in female speech.

Another unexpected discovery in this study is the finding that stroke count did not change significantly across DAF conditions. A study by Brendel et al. (2004) investigated the effects of DAF and frequency-shifted feedback (FSF) on various speech parameters, particularly focusing on speakers with hypokinetic dysarthria, often associated with Parkinson's disease (PD). The primary aim was to understand how these altered feedback conditions impact speech rate, intelligibility, naturalness, and other prosodic features. The study found that all speaker groups, including those with PD, showed a significant reduction in speaking rate under DAF. This reduction was mainly achieved by lengthening speech sounds and increasing the frequency of shorter pauses. Based on previous research, it would be reasonable to expect higher stroke counts if DAF conditions led to slower speech. Adams et al. (1993) found that slower speaking rates were associated with more velocity peaks in jaw movements. Although the count of velocity peaks is computed differently from the stroke count, these measures are equivalent in reflecting a more halting pattern of articulatory movements.

Anecdotally, many speakers in the present study seemed less fluent when subjected to DAF during the experiment, exhibiting hesitations, repetitions, and prolongations in their speech. This is consistent with the disrupted fluency caused by delays in auditory feedback reported by Chon et al. (2021). These disruptions in fluency were expected to increase the number of articulatory strokes, as speakers might adjust their speech motor patterns in response to the auditory feedback delays. However, the consistent stroke count across conditions suggests that while DAF impacts some aspects of speech, it does not necessarily alter the velocity profile of articulatory movements.

The findings of Nissen et al. (2007) show that higher accent ratings in L2 correlated with lower stroke speeds and speech-to-pause ratios highlight how kinematic differences contribute to perceptual aspects of speech. The current study did not directly measure perceptual features but did reveal that DAF-induced changes in peak speed and stroke distance could potentially disrupt fluent speech production, similar to how slower stroke speeds in L2 may affect accent perception. Future research could explore the intersection of DAF, speech kinematics, and perceived fluency, further linking the mechanical aspects of speech to listener perceptions.

Both studies underscore the relationship between articulatory precision, effort, and speech conditions. The Nissen et al. (2007) study suggests that slower L2 stroke speeds reflect increased articulatory effort as speakers lacked native automaticity. The current findings similarly show that DAF impacts peak speeds and stroke distances, likely increasing cognitive load and reducing the ease of articulation. This aligns with theories of speech production that propose greater effort and reduced automaticity under challenging conditions, whether due to speaking in an L2 or under DAF.

Limitations of the Present Study

One limitation in this study was the absence of perceptual ratings to assess how DAF influenced the perceived fluency and naturalness of speech. Without these ratings, it is challenging to associate the kinematic data with subjective listener impressions. Future studies could incorporate self-reports from participants to provide insight into their subjective experiences and perceptions of how DAF affected their speech. Additionally, the study used a smartphone app to administer DAF rather than a professional-grade system, which may have introduced some variability in feedback timing. The headphones used in the study, while functional, did not fully block the participants' own auditory feedback. Using higher-quality, noise-canceling headphones could have provided a more controlled auditory environment. These factors suggest that while the study's setup was effective, there are areas where improvements could enhance the reliability and precision of the findings.

Conclusions

This study investigated the effects of DAF on articulatory kinematics during tongue twisters and monologues and found that DAF significantly reduced peak speed and stroke distance for all three articulators during tongue twisters. Sex differences were evident, with women generally showing higher speeds and longer stroke distances. In monologues, DAF had no significant effects on peak speed, stroke distance, or hull area, but sex differences persisted, with men typically having lower stroke speeds, shorter stroke distances, and smaller hull areas.

The findings suggest that DAF disrupts motor performance and increases articulatory effort, particularly in structured and challenging tasks like tongue twisters. Sex differences in speech kinematics highlight the need to consider sex as a variable in speech production studies. These results have theoretical implications for understanding the role of altered feedback in speech motor control and clinical implications for developing targeted speech therapy interventions that consider sex differences and task-specific demands. Future research could explore the perceptual impact of DAF-induced changes in speech and further investigate the relationship between articulatory precision, effort, and listener perceptions of fluency.

References

Adams, S. G., Weismer, G., & Kent, R. D. (1993). Speaking rate and speech movement velocity profiles. *Journal of Speech and Hearing Research*, *36*(1), 41–54. https://doi.org/10.1044/jshr.3601.41

Brendel, B., Lowit, A., & Howell, P. (2004). The effects of delayed and frequency shifted feedback on speakers with Parkinson's Disease. *Journal of Medical Speech-Language Pathology, 12*(4)*,* 131–138.

Chon, H., Jackson, E. S., Kraft, S. J., Ambrose, N. G., & Loucks, T. M. (2021). Deficit or difference? Effects of altered auditory feedback on speech fluency and kinematic variability in adults who stutter. *Journal of Speech, Language, and Hearing Research*, *64*(7)*,* 2539–2556. https://doi.org/10.1044/2021_jslhr-20-00606

Dagenais, P. A., Southwood, M. H., & Lee, T. L. (1998). Rate reduction methods for improving speech intelligibility of dysarthric speakers with Parkinson's Disease. *Journal of Medical Speech-Language Pathology, 6*(3), 143–157.

Hain, T. C., Burnett, T. A., Kiran, S., Larson, C. R., Singh, S., & Kenney, M. K. (2000). Instructing subjects to make a voluntary response reveals the presence of two components to the audio-vocal reflex. *Experimental Brain Research, 130*(2), 133–141. https://doi.org/10.1007/s002219900237

Jones, J. A., & Striemer, D. (2007). Speech disruption during delayed auditory feedback with simultaneous visual feedback. *The Journal of the Acoustical Society of America*, *122*(4). https://doi.org/10.1121/1.2772402

Katseff, S., Houde, J., & Johnson, K. (2012). Partial compensation for altered auditory feedback: A tradeoff with somatosensory feedback? *Language and Speech*, *55*(2), 295–308. https://doi.org/10.1177/0023830911417802

Lincoln, M., Packman, A., & Onslow, M. (2006). Altered auditory feedback and the treatment of stuttering: A review. *Journal of Fluency Disorders*, *31*(2), 71–89. https://doi.org/10.1016/j.jfludis.2006.04.001

Mathworks. (2023). *Matlab* (Version 2023a) [Computer software]. https://www.mathworks.com/

Nissen, S. L., Dromey, C., & Wheeler, C. (2007). First and second language tongue movements in Spanish and Korean bilingual speakers. *Phonetica, 64*(4), 201–216. https://doi.org/10.1159/000121373

Northern Digital Inc. (2017). *Wavefront* (Version 2.2.1) [Computer software]. https://www.ndigital.com/

Svirsky, M. A., Lane, H., Perkell, J. S., & Wozniak, J. (1992). Effects of short-term auditory deprivation on speech production in adult cochlear implant users. *The Journal of the Acoustical Society of America*, *92*(3), 1284–1300. https://doi.org/10.1121/1.403923

Tasko, S. M., & Westbury, J. R. (2002). Defining and measuring speech movement events. *Journal of Speech, Language, and Hearing Research*, *45*(1), 127–142. https://doi.org/10.1044/1092-4388(2002/010)

Taylor, S., Dromey, C., Nissen, S. L., Tanner, K., Eggett, D., & Corbin-Lewis, K. (2020). Age-related changes in speech and voice: Spectral and cepstral measures. *Journal of Speech, Language, and Hearing Research*, *63*(3), 647–660. https://doi.org/10.1044/2019_JSLHR-19-00028

Te, G. O., Hamilton, M. J., Rizer, F. M., Schatz, K. A., Arkis, P. N., & Rose, H. C. (1996). Early speech changes in children with multichannel cochlear implants. *Otolaryngology—Head and Neck Surgery: Official Journal of American Academy of Otolaryngology-Head and Neck Surgery, 115*(6), 508–512. https://doi.org/10.1016/S0194-59989670004-9

Tables

Table 1

Means and Standard Errors for the Speech Kinematic Measures for the Tongue Twister and Monologue Tasks as a Function of DAF

	Tongue Twister								Monologue			
	No DAF		50 ms DAF		100 ms DAF		150 ms DAF		No DAF		150 ms DAF	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
TF Stroke Count	30.8	1.48	31.2	1.48	30.5	1.47	30.9	1.48	281.0	12.75	264.7	12.75
TF Peak Speed	66.6	2.48	59.9	2.46	60.6	2.45	59.6	2.46	97.2	3.84	95.3	3.84
TF Stroke Distance	6.8	0.25	6.4	0.24	6.4	0.24	6.3	0.24	9.8	0.36	9.7	0.36
TF Hull Area	176.0	16.27	168.9	16.12	187.9	16.04	186.4	16.10	1209.0	182.25	1142.0	182.25
J Stroke Count	30.1	1.62	30.2	1.61	29.6	1.61	30.6	1.61	270.3	12.15	257.1	12.15
J Peak Speed	29.0	1.68	27.0	1.66	26.7	1.66	25.9	1.66	40.9	2.61	39.0	2.61
J Stroke Distance	3.2	0.23	3.0	0.23	3.0	0.23	2.9	0.23	4.5	0.30	4.3	0.30
J Hull Area	46.7	8.22	48.0	8.12	52.4	8.06	52.1	8.11	911.2	166.20	836.9	166.20
LL Stroke Count	30.4	1.56	31.9	1.55	31.0	1.54	31.5	1.54	281.1	12.90	269.0	12.90
LL Peak Speed	34.3	1.61	32.0	1.59	31.9	1.59	31.1	1.59	59.7	3.21	57.1	3.21
LL Stroke Distance	3.8	0.22	3.6	0.22	3.6	0.22	3.5	0.22	6.2	0.34	5.9	0.34
LL Hull Area	65.6	8.89	68.8	8.75	72.3	8.68	73.4	8.74	1034.0	177.05	994.0	177.05

Note. DAF = delayed auditory feedback, SE = standard error, TF = tongue front, J = jaw, LL = lower lip. Speed in m per second. Distance in millimeters. Hull area mm2.

Table 2

Means and Standard Errors for the Speech Kinematic Measures for the Tongue Twister and Monologue Tasks for Men and Women

	Tongue Twister				Monologue			
	Male		Female		Male		Female	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE
TF Stroke Count	29.5	1.96	32.2	2.08	342.6	15.81	203.0	14.79
TF Peak Speed	65.1	3.20	58.3	3.39	94.7	4.98	97.8	4.66
TF Stroke Distance	6.6	0.31	6.4	0.33	9.4	0.48	10.1	0.45
TF Hull Area	162.3	20.95	197.3	22.19	843.5	249.32	1507.4	233.22
J Stroke Count	29.4	2.13	30.8	2.26	330.9	14.99	196.5	14.02
J Peak Speed	23.4	2.18	30.9	2.31	33.7	3.46	46.2	3.24
J Stroke Distance	2.4	0.30	3.6	0.32	3.6	0.40	5.2	0.37
J Hull Area	30.7	10.30	68.9	10.91	557.1	213.15	1191.0	199.39
LL Stroke Count	30.8	2.03	31.5	2.16	345.1	15.73	204.9	14.71
LL Peak Speed	29.6	2.08	35.1	2.21	50.0	4.35	66.7	4.06
LL Stroke Distance	3.1	0.28	4.2	0.30	5.1	0.46	7.0	0.43
LL Hull Area	49.7	10.91	90.4	11.55	653.7	224.72	1374.2	210.20

Note. SE = standard error, TF = tongue front, J = jaw, LL = lower lip.

Table 3

Linear Mixed Model ANOVA Results for the Tongue Twister Task

	DAF				Speaker Sex			
	df num	df den	*F*	*p*	df num	df den	*F*	*p*
TF Stroke Count	3	227.17	0.399	.754	1	15.01	0.935	.349
TF Peak Speed	3	227.31	12.197	< .001	1	15.04	2.091	.169
TF Stroke Distance	3	227.39	5.691	< .001	1	15.05	0.276	.607
TF Hull Area	3	227.30	2.272	.081	1	15.02	1.317	.269
J Stroke Count	3	227.18	0.766	.514	1	15.00	0.225	.642
J Peak Speed	3	227.23	4.893	.003	1	15.01	5.627	.031
J Stroke Distance	3	227.19	2.707	.046	1	15.00	7.724	.014
J Hull Area	3	227.35	0.620	.603	1	14.94	6.501	.022
LL Stroke Count	3	227.18	1.620	.185	1	14.98	0.057	.814
LL Peak Speed	3	227.17	5.335	< .001	1	14.92	3.267	.091
LL Stroke Distance	3	227.17	3.680	.013	1	14.96	6.933	.019
LL Hull Area	3	233.52	0.655	.581	1	19.00	6.581	.019

Note. TF = tongue front, J = jaw, LL = lower lip.

Table 4

Linear Mixed Model ANOVA Results for the Monologue Task

	DAF				Speaker Sex			
	df num	df den	*F*	*p*	df num	df den	*F*	*p*
TF Stroke Count	1	14.00	1.469	.246	1	13.00	41.587	< .001
TF Peak Speed	1	14.00	0.307	.588	1	13.00	0.200	.662
TF Stroke Distance	1	14.00	0.285	.602	1	13.00	0.993	.337
TF Hull Area	1	14.00	0.275	.608	1	13.00	3.782	.074
J Stroke Count	1	14.00	1.020	.330	1	13.00	42.877	< .001
J Peak Speed	1	14.00	0.759	.398	1	13.00	6.952	.021
J Stroke Distance	1	14.00	1.024	.329	1	13.00	8.952	.010
J Hull Area	1	14.00	0.218	.648	1	13.00	4.717	.049
LL Stroke Count	1	14.00	0.724	.409	1	13.00	42.369	< .001
LL Peak Speed	1	14.01	1.161	.299	1	13.00	7.933	.015
LL Stroke Distance	1	14.00	1.520	.228	1	13.00	9.693	.008
LL Hull Area	1	14.00	0.052	.823	1	13.00	5.483	.036

Note. TF = tongue front, J = jaw, LL = lower lip.

Figures

Figure 3

Mean and Standard Error of Tongue Front Peak Speed in the Tongue Twister Task Across DAF Conditions

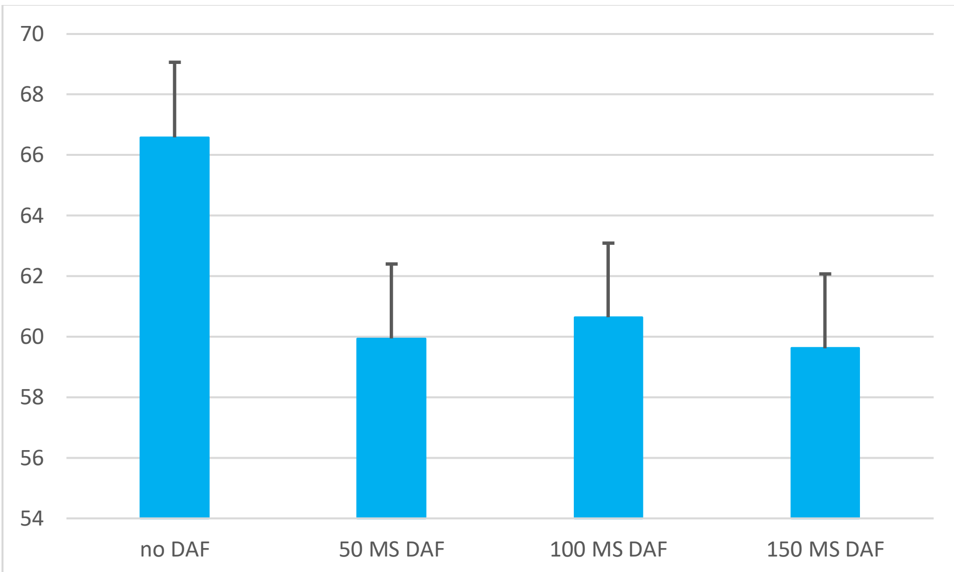


Figure 4

Mean and Standard Error of Jaw Stroke Distance in the Tongue Task

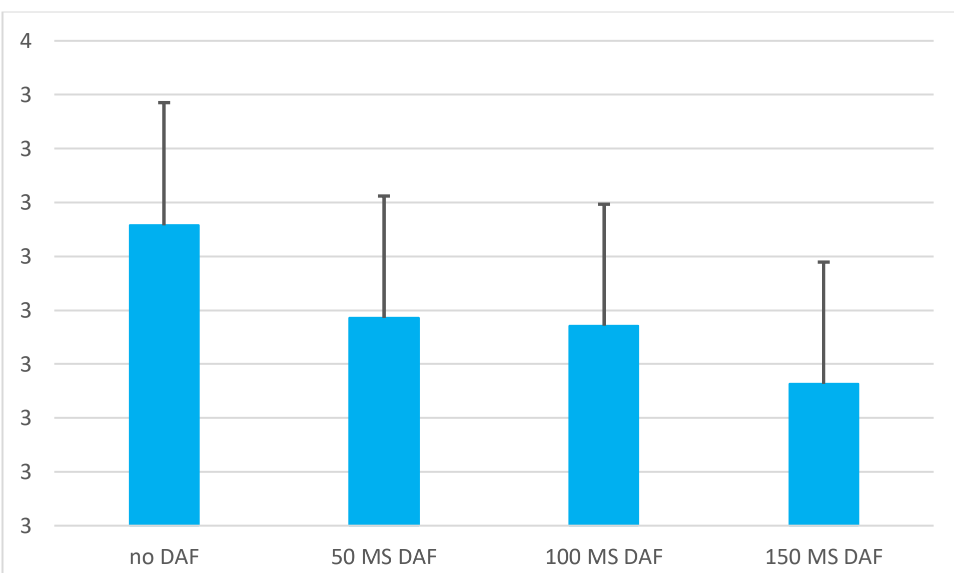


Figure 5

Mean and Standard Error of Lower Lip Hull Area in the Tongue Twister Task for Male and Female Speakers

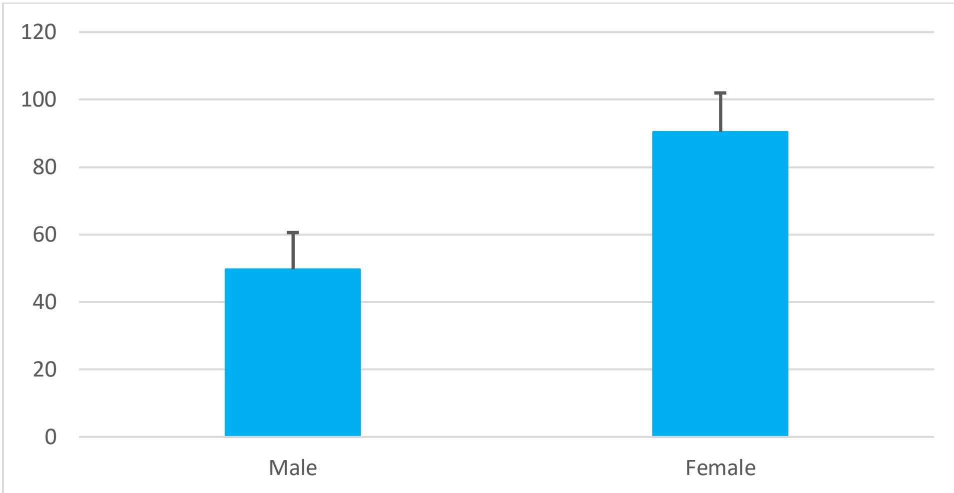


Figure 6

Mean and Standard Error of Tongue Front Stroke Count in the Monologue Task

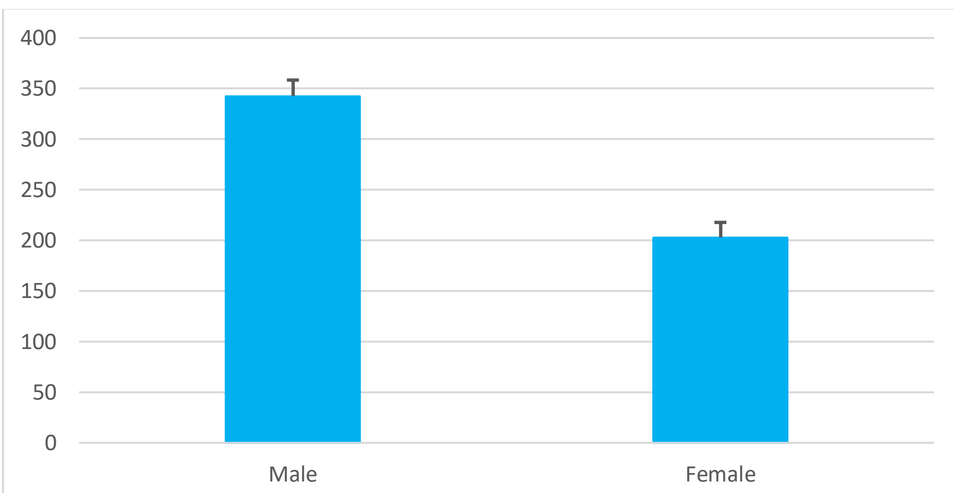


Figure 7

Mean and Standard Error of Jaw Peak Speed in the Monologue Task

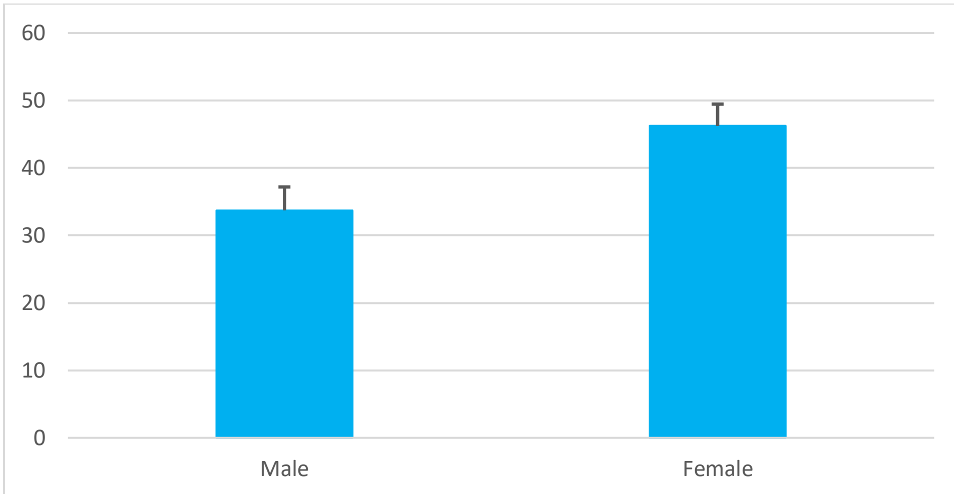
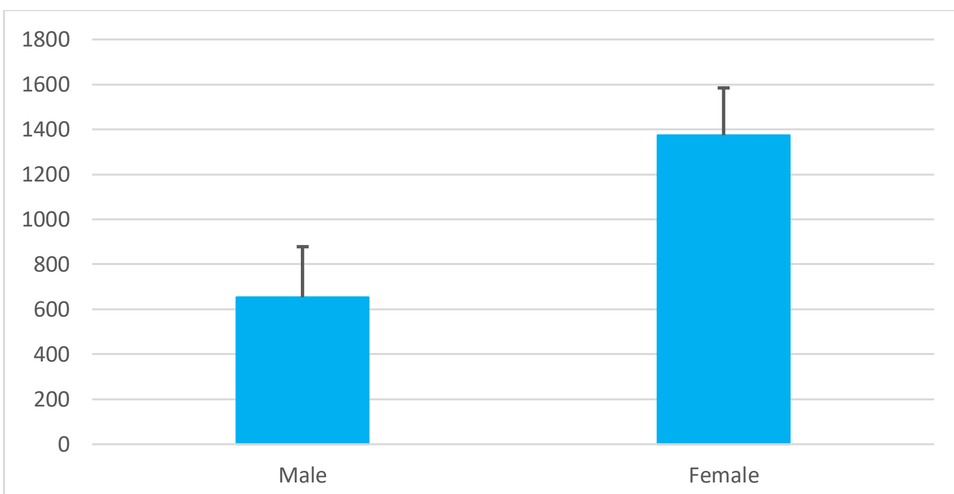


Figure 8

Mean and Standard Error of Lower Lip Hull Area in the Monologue Task



APPENDIX A

Annotated Bibliography

Adams, S. G., Weismer, G., & Kent, R. D. (1993). Speaking rate and speech movement velocity profiles. *Journal of Speech and Hearing Research*, *36*(1), 41–54. https://doi.org/10.1044/jshr.3601.41

Objective: To investigate the impact of speaking rate on the velocity profiles of lower lip and tongue tip movements during stop consonant production, using kinematic analysis.

Method: The study utilized an x-ray microbeam system to examine the velocity profiles of lower lip and tongue tip movements during stop consonant production. Five young adults participated in a magnitude production task to produce five different speaking rates, ranging from very fast to very slow.

Results: The findings indicated that speaking rate changes were associated with alterations in the topology of the speech movement velocity-time function. Fast speaking rates exhibited a symmetrical, single-peaked velocity profile, whereas slow speaking rates showed an asymmetrical, multi-peaked profile.

Conclusions: The study concluded that variations in speaking rate are linked to changes in motor control strategies. Fast speaking rates likely involve unitary, preprogrammed movements, while slow speaking rates consist of multiple submovements influenced by feedback mechanisms.

Relevance to Current Study: This research is relevant to studies using kinematics to study Delayed Auditory Feedback (DAF) as it highlights how motor control strategies in speech production vary with speaking rate, providing insights into how DAF might influence speech motor control and feedback mechanisms.

Brendel, B., Lowit, A., & Howell, P. (2004). The effects of delayed and frequency shifted feedback on speakers with Parkinson's Disease. *Journal of Medical Speech-Language Pathology*, *12*(4), 131–138.

Objective: To investigate the effects of DAF and FSF on the speech of Parkinson's Disease (PD) patients and control participants across a broad range of speech measures.

Method: This study assessed the effects of DAF and FSF on the speech of 16 speakers with Parkinson's Disease (PD) and 11 control speakers. Participants performed a reading task under three conditions: DAF, FSF, and no altered feedback (NAF).

Results: Both altered feedback conditions resulted in a significant reduction in speech rate across all groups. Additional changes included increases in pause frequency, loudness levels, and pitch variation, while intelligibility and naturalness decreased for all or some groups. Minimal effects were observed on articulation/pause time ratio, pause duration, pitch range, and speech rhythm. FSF generally resulted in performance closer to the NAF state than DAF.

Conclusions: The study confirmed that PD speakers respond differently to altered feedback, with some benefiting from the system despite overall negative effects on intelligibility and naturalness. FSF was found to produce more natural speech compared to DAF.

Relevance to Current Study: This research is relevant to understanding how altered auditory feedback, such as DAF and FSF, influences speech in PD patients. It provides insights into the effectiveness of these tools for speech therapy, highlighting the nuanced responses of PD speakers and the potential benefits of FSF over DAF.

Chon, H., Jackson, E. S., Kraft, S. J., Ambrose, N. G., & Loucks, T. M. (2021). Deficit or difference? Effects of altered auditory feedback on speech fluency and kinematic variability in adults who stutter. *Journal of Speech, Language, and Hearing Research*, *64*(7), 2539–2556. https://doi.org/10.1044/2021_jslhr-20-00606

Objective: This experiment tested whether adults who stutter (AWS) would display a different range of sensitivity to delayed auditory feedback (DAF) from adults who do not stutter (AWNS).

Method: In Experiment 1, 15 AWS performed a conversational speaking task under non altered auditory feedback and 250-ms DAF. The rates of stuttering-like disfluencies, other disfluencies, speech errors, and articulation rate were compared. In Experiment 2, 13 AWS and 15 AWNS read three utterances under four auditory feedback conditions: nonaltered auditory feedback, amplified auditory feedback, 25-ms DAF, and 50-ms DAF. Across-utterance kinematic variability and within-utterance variability were compared between groups.

Results: The rate of stuttering-like disfluencies and speech errors increased significantly, while articulation rate decreased significantly in AWS under 250-ms DAF. AWS exhibited higher kinematic variability than AWNS across the feedback conditions.

Conclusions: Auditory feedback manipulations can alter speech fluency and kinematic variability in adults who stutter.

Relevance to Current Study: Testing the effects of different latencies of DAF is the objective of the current study. Therefore, the information in this article is relevant to the current study as it provides information about different latencies of DAF and its clinical applications.

Cler, G. J., Lee, J. C., Mittelman, T., Stepp, C. E., & Bohland, J. W. (2017). Kinematic analysis of speech sound sequencing errors induced by delayed auditory feedback. *Journal of Speech, Language, and Hearing Research*, *60*(6S), 1695–1711. https://doi.org/10.1044/2017_jslhr-s-16-0234

Objective: This study collected kinematic data to assess the articulatory features of disfluencies and phonological errors as a result of DAF.

Method: Eight typical speakers produced nonsense syllable sequences under normal and DAF (200 ms) conditions. Lip and tongue kinematics were captured with electromagnetic articulography. Time-locked acoustic recordings were transcribed, and the kinematics of utterances with and without perceived errors were analyzed with existing and novel quantitative methods.

Results: For five participants, kinematic variability for productions perceived to be error-free significantly increased under DAF.

Conclusions: This research represents one of the initial efforts to describe articulatory alterations with Delayed Auditory Feedback (DAF) and offers supporting proof for distinct categories of speech errors that might not be easily noticeable. Novel techniques were created to facilitate the visualization and analysis of extensive kinematic data collections.

Relevance to Current Study: The current study measures the effects of DAF using kinematics, so the information in this article is relevant to the current study as it reports on speech kinematic measurement.

Dagenais, P. A., Southwood, M. H., & Lee, T. L. (1998). Rate reduction methods for improving speech intelligibility of dysarthric speakers with Parkinson's Disease. *Journal of Medical Speech-Language Pathology, 6*(3), 143–157.

Methods: Three speakers with moderate idiopathic Parkinson's disease and dysarthria participated in the study. Each speaker underwent a speech therapy program with varying protocols, including delayed auditory feedback (DAF), DAF plus traditional clinician-directed therapy, and DAF plus clinician-directed prolonged speech therapy. A single-subject multiple baseline design was used to assess changes in speech intelligibility during reading, picture description, and spontaneous speech tasks.

Results: The results showed that continued exposure to DAF did not lead to reductions in speaking rate. However, combining DAF with prolonged speech therapy had varying effects on the speakers. Speaker 1 demonstrated a reduction in speech rate and improvement in intelligibility scores, which were maintained even when the speaking rate was gradually increased. Speaker 2 showed erratic speech rate and intelligibility scores across all therapy protocols. Speaker 3 did not respond to the treatment protocols but did show an increase in post-therapy intelligibility scores compared to pre-therapy scores.

Conclusions: The study suggests that individuals with Parkinson's disease may have cognitive deficits affecting monitoring skills and attention, which can impact speech intelligibility. The effectiveness of different therapy protocols varied among the speakers, highlighting the need for individualized treatment approaches.

Relevance to the Study: This study provides valuable insights into the effectiveness of speech therapy programs for individuals with Parkinson's disease.

Understanding the impact of speech rate reduction on speech intelligibility can help in developing more targeted and effective treatment strategies for individuals with Parkinson's disease and dysarthria.

Dromey, C., & Black, K. M. (2017). Effects of laryngeal activity on articulation. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, *25*(12), 2272–2280. https://doi.org/10.1109/taslp.2017.2738564

Objective: This study examined the effects of three speech conditions (voiced, whispered, and mouthed) on global average measures of articulatory movement during sentence production.

Method: Participants were 20 adults who produced six target utterances in the three speaking conditions. Movements of articulators were recorded with an electromagnetic articulograph.

Results: Measures revealed a number of significant changes between the voiced and mouthed conditions, with relatively few differences between voiced and whispered speech.

Conclusions: These findings suggest that both laryngeal activation and auditory feedback play an important role in the production of normally articulated vocal tract movements, and that the absence of these may account for the significant changes in articulation between the voiced and mouthed conditions.

Relevance to Current Study: The current study used sensors attached to the participant's tongue, lips, and jaw to measure speech movements. Therefore, the information in this article is relevant to the current study as it provides information about the effects of sensors on speech production in different conditions.

Dromey, C., Hunter, E., & Nissen, S. L. (2018). Speech adaptation to kinematic recording sensors: Perceptual and acoustic findings. *Journal of Speech, Language, and Hearing Research*, *61*(3), 593–603. https://doi.org/10.1044/2017_jslhr-s-17-0169

Objective: This study used perceptual and acoustic measures to examine the time course of speech adaptation after the attachment of electromagnetic sensor coils to the tongue, lips, and jaw.

Method: Twenty native English speakers read aloud stimulus sentences before the attachment of the sensors, immediately after attachment, and again 5, 10, 15, and 20 min later. Sentence recordings were perceptually evaluated by 20 native English listeners, who rated 150 stimuli using a visual analog scale with the end points labeled as “precise” and “imprecise.”

Results: Perceptual ratings revealed a decrease in speech precision after sensor attachment and evidence of adaptation over time; there was little perceptual change beyond the 10-min recording.

Conclusions: The results indicate that a 10-minute period might offer adequate time for speakers to acclimate before commencing the experimental data collection using Northern Digital Instruments Wave electromagnetic sensors.

Relevance to Current Study: The current study used sensors attached to the participant’s tongue, lips, and jaw to measure speech movements. Therefore, the information in this article is relevant to the current study as it provides information about the effects of sensors on speech production.

Hain, T. C., Burnett, T. A., Kiran, S., Larson, C. R., Singh, S., & Kenney, M. K. (2000). Instructing subjects to make a voluntary response reveals the presence of two components to the audio-vocal reflex. *Experimental Brain Research, 130*(2), 133–141. https://doi.org/10.1007/s002219900237

Objective: This study aimed to investigate the hypothesis that subjects exhibit two distinct vocal responses to pitch-shift stimuli and to clarify the influence of intention on these responses.

Methods: Subjects were instructed to modify their voice fundamental frequency (F0) in response to pitch shifts presented to them. They were instructed to either change their voice F0 in the opposite direction of the pitch shift, in the same direction, or not to respond at all. Another group of subjects was instructed to raise or lower their voice F0 as rapidly as possible upon hearing a pitch shift. The study measured the subjects' vocal responses (VR1 and VR2) under these different instructions.

Results: The results showed that subjects, when instructed to produce a voluntary response, exhibited both an early vocal response (VR1) and a later vocal response (VR2) to the pitch-shift stimuli. VR2 was almost always made in the instructed direction, while VR1 was often incorrect. The latency of VR1 was reduced when subjects were instructed to change voice F0 in the opposite direction of the pitch shift compared to when they were told to ignore the pitch shifts. Latency and amplitude measures of VR2 varied under different experimental conditions.

Conclusions: The study demonstrated that subjects exhibit two distinct vocal responses to pitch-shift stimuli. The first response (VR1) appears to be relatively automatic but can be modulated by instructions to the participant. The second response (VR2) is likely a voluntary one. These findings suggest that the modulation of VR1 and the production of VR2 are influenced by the intention and instructions given to the subjects.

Relevance to Current Study: This study provides insights into the mechanisms underlying vocal responses to pitch-shift stimuli. Understanding how intention and instructions influence these responses can have implications for studies on speech motor control and may contribute to the development of strategies for improving speech therapy for individuals with speech disorders.

Jones, J. A., & Striemer, D. (2007). Speech disruption during delayed auditory feedback with simultaneous visual feedback. *The Journal of the Acoustical Society of America*, *122*(4), EL135–EL141. https://doi.org/10.1121/1.2772402

Objective: This study explored whether providing visual feedback in addition to DAF would negatively affect speech.

Method: Participants sat in a double-walled sound booth and wore headphones and a headset microphone while their vocal productions were recorded as they repeated the same ten sentences in each experimental condition.

Results: Sentence durations in the DAF conditions were longer than sentence durations in the NAF conditions and there was a greater number of disruptions in the DAF conditions than in the NAF conditions.

Conclusions: Visual feedback does not mitigate the impact of DAF. Nevertheless, the observations indicate encouraging patterns that imply visual feedback may have potential in this regard.

Relevance to Current Study: One of the speaking conditions in the current study is delayed auditory feedback (DAF). Therefore, the information in this article is relevant to the current study as it provides information about DAF and its clinical applications.

Katseff, S., Houde, J., & Johnson, K. (2012). Partial compensation for altered auditory feedback: A tradeoff with somatosensory feedback? *Language and Speech, 55*(2), 295–308. https://doi.org/10.1177/0023830911417802

Objective: This study aimed to characterize and understand partial compensation in speech production by examining how talkers respond to each step on a staircase of increasing shifts in auditory feedback.

Method: Subjects wore an apparatus that altered their real-time auditory feedback. They were asked to repeat visually-presented hVd stimulus words while feedback was altered stepwise over the course of 360 trials. A novel analysis method was used to calculate each subject's compensation at each step relative to their baseline.

Results: The results demonstrated that subjects compensated more for small feedback shifts than for larger shifts. This pattern suggests that vowel targets incorporate both auditory and somatosensory information, and the speech motor control system is driven by differential weighting of auditory and somatosensory feedback.

Conclusion: The findings provide insight into the mechanisms underlying speech production and the role of auditory and somatosensory feedback in shaping vowel targets. The differential compensation for small and large feedback shifts highlights the complex interplay between sensory feedback and motor control in speech.

Relevance to current study: This study's investigation into partial compensation for experimentally-induced changes in auditory feedback provides valuable insights for the current study on delayed auditory feedback (DAF) and its effects on speech production. The findings regarding how talkers respond to different levels of auditory feedback alteration can inform the understanding of how individuals might adapt to the altered auditory feedback experienced in DAF conditions. Additionally, the study's novel analysis method for calculating compensation at each step relative to baseline can offer a model for analyzing the effects of DAF on speech using stroke measures, providing a framework for assessing compensatory mechanisms in speech motor control.

Lincoln, M., Packman, A., & Onslow, M. (2006). Altered auditory feedback and the treatment of stuttering: A review. *Journal of Fluency Disorders*, *31*(2), 71–89. https://doi.org/10.1016/j.jfludis.2006.04.001

Objective: This paper reviews published, peer-reviewed journal papers from the past 10 years that investigate the effect of altered auditory feedback (AAF) devices as a treatment for adults and children who stutter, aiming to assess the current state of knowledge regarding the effectiveness of AAF in different speaking conditions, tasks, and situations.

Method: The review examines experimental evidence and limited Phase 1 treatment outcome evidence regarding the effect of AAF on the speech of people who stutter. It also considers gaps in the literature, such as the lack of knowledge about the effect of AAF during conversational speech and in everyday speaking situations, as well as the need to establish how to determine the correct levels of AAF for individuals and the characteristics of those likely to benefit from AAF.

Results: The review indicates that while considerable experimental evidence has been accumulated regarding the effect of AAF on stuttering, critical knowledge gaps remain. Specifically, there is insufficient evidence to support the use of AAF devices as a clinical option for children who stutter. Additionally, the development and availability of AAF devices have outpaced clinical trials research.

Conclusion: Despite the advancements in AAF device development, there is still a need for further research to determine the effectiveness and appropriate use of AAF in treating stuttering, especially in real-world speaking situations. Future research should focus on addressing these knowledge gaps to better inform the clinical use of AAF devices for individuals who stutter.

Relevance to Current Study: The review of literature on altered auditory feedback (AAF) devices in stuttering treatment offers relevant insights for the current study on delayed auditory feedback (DAF) and its effects on speech production. While AAF and DAF differ in their mechanisms, both involve altering auditory feedback to potentially improve speech fluency. The review's discussion on the need for determining the correct levels of AAF for individuals and identifying characteristics of those likely to benefit from AAF can inform similar considerations in designing studies using DAF with stroke measures. Furthermore, the review's emphasis on the gaps in knowledge about AAF's effects during conversational speech and in everyday speaking situations underscores the importance of examining DAF's impact beyond controlled laboratory settings in the current study.

Malloy, J. R., Nistal, D., Heyne, M., Tardif, M. C., & Bohland, J. W. (2022). Delayed auditory feedback elicits specific patterns of serial order errors in a paced syllable sequence production task. *Journal of Speech, Language, and Hearing Research*, *65*(5), 1800–1821. https://doi.org/10.1044/2022_jslhr-21-00427

Objective: The goal of this study was to characterize the types of serial order errors that increase under DAF in a systematic syllable sequence production task, which used a closed set of sounds and controlled for speech rate.

Method: Sixteen adult speakers repeatedly produced CVCVCV sequences, paced to a “visual metronome,” while hearing self-generated feedback with delays of 0–250 ms. Listeners transcribed recordings, and speech errors were classified based on the literature about naturally occurring slips of the tongue. A series of mixed-effects models were used to assess the effects of delay for different error types, for error arrival time, and for speaking rate.

Results: DAF had a significant effect on the overall error rate for delays of 100 ms or greater. Statistical models revealed significant effects for vowel and syllable repetitions, vowel exchanges, vowel omissions, onset disfluencies, and distortions. Serial order errors were especially dominated by vowel and syllable repetitions. Errors occurred earlier on average within a trial for longer feedback delays.

Conclusions: DAF induces a distinct sequence of serial order errors. The primary sequence of errors, characterized by vowel and syllable repetition, hints at potential mechanisms through which DAF alters the functioning of speech planning representations, resulting in errors.

Relevance to Current Study: One of the speaking conditions in the current study is delayed auditory feedback (DAF). Therefore, the information in this article is relevant to the current study as it provides information about DAF and its clinical applications.

Nissen, S. L., Dromey, C., & Wheeler, C. (2007). First and second language tongue movements in Spanish and Korean bilingual speakers. *Phonetica, 64*(4), 201–216. https://doi.org/10.1159/000121373

Objective: To investigate intraspeaker differences in the production of native (Korean or Spanish) and second language (English) using kinematic indices of tongue activity.

Method: The study investigated intraspeaker differences in native (Korean or Spanish) and second language (English) production using kinematic indices of tongue activity. Measurements included the speed, duration, and distance of tongue movements during speech.

Results: Speakers exhibited significantly slower stroke speeds and longer movement durations for their L2 compared to their L1, with no significant differences in stroke distance. Bilingual speakers paused more and spoke less in their L2. Those with greater kinematic changes from L1 to L2 were perceived to have a stronger accent.

Conclusions: The findings suggest that bilingual speakers display different articulatory behaviors in their L2, with slower and longer tongue movements and increased pausing, which contribute to a stronger perceived accent.

Relevance to Current Study: This research highlights the importance of kinematic analysis in understanding L2 production. It provides valuable insights into the physical articulatory differences between L1 and L2 speech, offering a more comprehensive understanding of second language acquisition and pronunciation.

Pouplier, M. (2007). Tongue kinematics during utterances elicited with the SLIP technique. *Language and Speech, 50*(3), 311–341. https://doi.org/10.1177/00238309070500030201

Objective: This study aims to investigate the nature of speech production errors by analyzing tongue movement data collected during an error elicitation study based on the SLIP technique. Specifically, the study seeks to determine whether the errors observed in tongue kinematics during the SLIP task are indicative of cognitive planning errors at the phonological level or are merely articulation errors.

Method: Tongue movement data were collected using the Perkell system electromagnetic midsagittal articulometer. Transducers were attached to the nose ridge, upper lip and lower lip, lower incisor, maxilla, and the tongue.

Results: Tongue kinematics during errors in the present task were comparable to those found in errorful utterances in repetition tasks

Conclusions: Error elicitation through SLIP produced the same type of errors as tongue-twister type tasks.

Relevance to Current Study: The current study used a tongue-twister as a speech stimulus, similar to the cited study.

Svirsky, M. A., Lane, H., Perkell, J. S., & Wozniak, J. (1992). Effects of short-term auditory deprivation on speech production in adult cochlear implant users. *The Journal of the Acoustical Society of America*, *92*(3), 1284–1300. https://doi.org/10.1121/1.403923

Objective: The objective of this study was to measure and analyze the speech production parameters of three postlingually deafened adults who use cochlear implants under different auditory conditions. Specifically, the study aimed to observe the changes in vowel acoustics and average airflow in these individuals after 24 hours of auditory deprivation, upon reactivation of their speech processors, and subsequent auditory deprivation.

Method: Three postlingually deafened adults who use cochlear implants participated in this study. The experiment consisted of three phases: Auditory Deprivation: The participants' speech processors were turned off for 24 hours. Reactivation: The speech processors were turned back on. Subsequent Deprivation: The speech processors were turned off again. During each phase, the following speech production parameters were measured while the participants read word lists and passages:

Results: Changes in the state of the speech processor (on-to-off or vice versa) were accompanied by numerous alterations in speech production parameters. Key findings include: Many changes in speech parameters were in the direction of normalcy. Most changes were consistent with long-term speech production changes observed in the same subjects following the activation of their cochlear implant processors, as reported by Perkell et al. (1992). Changes in mean airflow were always accompanied by H1–H2 (breathiness) changes in the same direction, suggesting underlying changes in laryngeal posture.

Conclusion: The findings of this study suggest that auditory feedback plays a dual role in speech production for postlingually deafened adults using cochlear implants. Auditory feedback contributes to the long-term calibration of articulatory parameters and supports feedback mechanisms with relatively short time constants. The observed rapid changes in speech production parameters upon reactivation of the speech processors highlight the immediate impact of auditory feedback, whereas the more gradual changes upon deactivation suggest an ongoing adjustment process in the absence of auditory input.

Relevance to Current Study: The current study aims to understand the role of auditory feedback on speech.

Tasko, S. M., & Westbury, J. R. (2002). Defining and measuring speech movement events. *Journal of Speech, Language, and Hearing Research, 45*(1), 127–142. https://doi.org/10.1044/1092-4388(2002/010)

Objective: Describe an approach for parsing kinematic signal streams into movement units.

Method: Speech materials recorded from 18 speakers were selected for analysis from the University of Wisconsin X-ray Microbeam Speech Production Database. Speakers performed an oral reading of a slightly expanded version of the Hunter script (Crystal & House, 1982) read at a self-selected speaking rate. Data acquisition was performed using the University of Wisconsin X-ray Microbeam system, according to the procedures described in Westbury (1994). Briefly, articulator motion was recorded by tracking the midsagittal positions of small gold pellets.

Results: There were typically fewer strokes than adjacent sound pairs, more strokes than syllables or words, and different numbers of strokes per articulator. About half of the observations fall between 95 and 172 ms for stroke duration. Half of the observed strokes had distances between 1.5 and 6.8 mm. Half of the values fell between 17 and 72 mm/s for peak stroke speed. About half the stroke boundary speeds fall between 3 and 15 mm/s.

Conclusions: There are some distinct advantages associated with this parsing method; for example, its strokes can be defined without reference to any terms external to the geometry of bodily movements. Thus, the approach can be applied to all types of motor tasks and is essentially independent of any assumptions about goals associated with such tasks.

Relevance to Current Study: The current study used stoke measures to reflect the kinematic features of the speech stimuli.

Taylor, S., Dromey, C., Nissen, S. L., Tanner, K., Eggett, D., & Corbin-Lewis, K. (2020). Age-related changes in speech and voice: Spectral and cepstral measures. *Journal of Speech, Language, and Hearing Research*, *63*(3), 647–660. https://doi.org/10.1044/2019_JSLHR-19-00028

Objective: To examine differences in selected acoustic measures of speech and voice according to age, sex, and familial relationships.

Method: The study examined acoustic measures of speech and voice in 169 participants (79 men and 90 women) from 18 families, aged 17 to 87 years. Participants read aloud two passages, and measures included fricative spectral moments (center of gravity, standard deviation, skewness, and kurtosis), proportion of time spent speaking, mean speaking fundamental frequency, semitone standard deviation (STSD), and cepstral peak prominence smoothed.

Results: Significant age effects were found for fricative spectral center of gravity, spectral skewness, and speaking STSD. Significant sex effects were observed for spectral center of gravity, spectral kurtosis, and mean fundamental frequency. Familial relationships significantly influenced spectral skewness, STSD, and cepstral peak prominence smoothed.

Conclusions: The study revealed that certain speech and voice features change with age and vary between men and women. Additionally, family members demonstrated similar patterns in prosody, voicing, and articulatory behavior. These findings highlight normal variations in speech and voice across age, sex, and family, which are important for distinguishing between normal and disordered patterns clinically.

Relevance to Current Study: This research provides insights into how demographic variables such as age, sex, and familial relationships influence speech and voice characteristics. Understanding these patterns is crucial for accurate clinical assessment and differentiation of normal and disordered speech and voice patterns.

Te, G. O., Hamilton, M. J., Rizer, F. M., Schatz, K. A., Arkis, P. N., & Rose, H. C. (1996). Early speech changes in children with multichannel cochlear implants. *Otolaryngology—Head and Neck Surgery: Official Journal of American Academy of Otolaryngology-Head and Neck Surgery, 115*(6), 508–512. https://doi.org/10.1016/S0194-59989670004-9

Objective: The study aimed to explore the early development of speech perception and production skills in prelingually deaf children who underwent cochlear implantation.

Methods: Four prelingually deaf children who received cochlear implants within a month of each other were observed for a year, with a focus on the first few months of rehabilitation. Speech scores were assessed immediately after the implant tune-up and monitored over the following months. Speech perception tests were conducted one year post-implantation.

Results: Immediate improvements in speech scores were observed, with continued rapid progress in speech production over the first four months. Progress slowed down after four months, especially in some speech production skills. Vowel production was the easiest skill to achieve, followed by word-pattern recognition and consonant voicing. Consonant placing and manner of consonant production were the most challenging skills.

Speech perception tests one year after implantation showed marked improvements in three out of four children compared to pre-implantation levels.

Conclusion: The study highlights the importance of early and intensive speech rehabilitation efforts for prelingually deaf children undergoing cochlear implantation to maximize the benefits of the implant.

Relevance to Current Study: The study on cochlear implants underscores the critical role of early and intensive speech rehabilitation in maximizing the benefits of auditory interventions for individuals with hearing impairments. Similarly, in the current study on Delayed Auditory Feedback (DAF), understanding the timing and effectiveness of speech skill development is crucial. By examining the early stages of speech perception and production in cochlear implant recipients, this prior research provides valuable insights into the potential trajectories and challenges of speech development under altered auditory conditions, such as those induced by DAF. Insights from the cochlear implant study can inform strategies for optimizing DAF-based interventions, especially in understanding the initial stages of skill acquisition and the long-term outcomes of DAF interventions for individuals with speech disorders

APPENDIX B

Consent Form

Consent to be a Research Subject

Title of the Research Study: Connecting lab speech with everyday communication Principal Investigator: Christopher Dromey, PhD
IRB ID#:

Introduction

This research study is being conducted by Professor Christopher Dromey, assisted by Lauren Clarke and Jessica Martin, all from the Department of Communication Disorders at Brigham Young University, to determine how speaking shorter or longer words and phrases affects how the tongue, lips, and jaw move. You were invited to participate because you are a native speaker of American English and have no history of speech, language, or hearing disorders.

Procedures

If you agree to participate in this research study, the following will occur:

- you will sit in a sound-treated booth in room 106 of the John Taylor Building where your speech will be recorded
- new, disposable small sensors (3 x 3 mm) will be attached with dental glue to your tongue, lips, and lower front teeth
- for 10 minutes you will either read aloud or chat with the researcher as you get used to the feeling of the sensors in your mouth
- you will read aloud words, phrases, and sentences that will be presented on a computer screen in front of you
- you will be asked to speak as you typically would and also with exaggerated clarity
- total time commitment will be 60 minutes in one recording session

Risks/Discomforts

You might feel uncomfortable having electromagnetic sensors attached to your tongue, lips, and lower teeth. It is possible that some of the dental glue will remain on the tongue surface for a few minutes after the experiment is over. This may feel odd, but it will feel normal again within a few minutes.

The researcher will view the surface of your tongue after removing the sensors to make sure that any traces of the glue are minimal. The single-use sensors will be thrown away after removal.

You might feel some fatigue; if so, you may take a break at any time during the study.

Benefits

There will be no direct benefits to you as a participant. However, we anticipate that the findings from this study will benefit the field of speech pathology by helping us design better treatments for people with speech problems.

Confidentiality

The research data will be kept on password protected computer and only the researchers will have access to the data. Before we analyze the recordings, all identifying information will be removed so that your name will not be linked to the recordings. Only summary data from groups of participants will be reported in publications and presentations. After the study the de-identified data will be kept on a password-protected computer in the researcher's office for possible future analysis with new techniques.

Compensation

You will receive $15 for your participation, whether you finish the recording or not; compensation will not be prorated.

Participation

Participation in this research study is voluntary. You have the right to withdraw at any time or refuse to participate entirely without jeopardy to your class status, grade, or standing with the university.

Questions about the Research

If you have questions regarding this study, you may contact Christopher Dromey at 133 TLRB, 801-422-6461, dromey@byu.edu for further information.

Questions about Your Rights as Research Participants

If you have questions regarding your rights as a research participant contact Human Research Protections Program by phone at (801) 422-1461; or by email: BYU.HRPP@byu.edu.

Statement of Consent

I have read, understood, and received a copy of the above consent and desire of my own free will to participate in this study.

Name (Printed): ____________________ Signature ______________________ Date: ________________________________

APPENDIX C

Institutional Review Board Approval

Memorandum

To: Christopher Dromey
Department: BYU - EDUC - Communications Disorders
From: Sandee Aina, MPA, HRPP Associate Director
Wayne Larsen, MAcc, IRB Administrator
Bob Ridge, Ph.D., IRB Chair
Date: December 13, 2022
IRB#: IRB2022-468
Title: Connecting lab speech with everyday communication

Brigham Young University's IRB has approved the research study referenced in the subject heading as expedited level, categories 4 and 6. The approval period is from 12/13/2022 to 12/12/2023. Thereafter, continued approval is contingent upon the submission of a continuing review request that must be reviewed and approved by the IRB prior to the expiration date of the study. Please reference your assigned IRB identification number in any correspondence with the IRB.

Continued approval is conditional upon your compliance with the following requirements:

1. A copy of the approved informed consent statement and associated recruiting documents (if applicable) can be accessed in iRIS. No other consent statement should be used. Each research subject must be offered a copy or provided a way to access the consent statement.
2. Any modifications to the approved protocol must be submitted, reviewed, and approved by the IRB before modifications are incorporated into the study.
3. All recruiting tools must be submitted and approved by the IRB prior to use.
4. All data, as well as the investigator's copies of the signed consent forms, must be retained for a period of at least three years following the termination of the study.
5. In addition, serious adverse events must be reported to the IRB immediately, with a written report by the PI within 24 hours of the PI's becoming aware of the event. Serious adverse events are (1) the death of a research participant; or (2) serious injury to a research participant.
6. All other non-serious unanticipated problems should be reported to the IRB within 2 weeks of the first awareness of the problem by the PI. Prompt reporting is important, as unanticipated problems often require some modification of study procedures, protocols, and/or informed consent processes. Such modifications require the review and approval of the IRB.

If it is necessary to continue the study beyond the expiration date, you will need to complete the continuing review form and attach associated documents to renew the study. Continuing review documents should be submitted no later than two months before 12/12/2023. More information regarding the renewal process and lapses in approval can be found on the IRB website FAQ #8.

There is no grace period beyond the expiration date. In order to avoid lapses in approval of your research and the possible suspension of subject enrollment, please look for notifications prompting you to initiate a continuing review request. You will receive two prompts from iRIS to renew this protocol, the IRB requires time to review your documents so **please be aware that requests made close to or on the expiration date will not be accepted.**