2024-06-04

Perceptual, Acoustic, and Kinematic Measures of Speech Precision and Steadiness

Jessica Jamiel Martin
Brigham Young University

Perceptual, Acoustic, and Kinematic Measures

of Speech Precision and Steadiness

Jessica Jamiel Martin

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
Tyson Harmon
Dallin Bailey

Department of Communication Disorders

Brigham Young University

ABSTRACT

Perceptual, Acoustic, and Kinematic Measures
of Speech Precision and Steadiness

Jessica Jamiel Martin
Department of Communication Disorders, BYU
Master of Science

Clinicians rely on perceptual analysis in the assessment and diagnosis of motor speech disorders. However, connecting perceptual measures to quantitative data has proved challenging. This study uses correlational analyses to explore the relationship between perceptual, acoustic, and kinematic measures. Twenty typical speakers provided speech samples of rapid syllable repetition and speech tasks, which were then rated by 12 listeners for precision and steadiness on a visual analog scale. Data was analyzed to identify significant correlations between the measures. We found evidence of a modest perceptual-acoustic relationship, with results suggesting that acoustic rate may be correlated with perceptual features. Our findings also suggest a significant perceptual-kinematic relationship, as several kinematic measures of displacement demonstrated significant correlations with precision and steadiness ratings. We found that speakers with more consistent speech movements received higher steadiness ratings, and speakers with faster articulatory movements were rated as more precise. This study supports the use of perceptual analysis in clinical practice and points towards establishing connections between perceptual, acoustic, and kinematic measures used in speech analysis.

Keywords: perceptual evaluation, speech acoustics, speech kinematics, rapid syllable repetition

ACKNOWLEDGMENTS

I would like to express my gratitude to Dr. Christopher Dromey, who was instrumental in the completion of this project. I am deeply indebted to him for his time, expertise, and mentorship over the last 2 years. I would also like to express my deepest appreciation to Dr. Tyson Harmon and Dr. Dallin Bailey for their invaluable feedback and support.

I am deeply grateful to the McKay School of Education for providing the funding that allowed us to complete this research. I would also like to extend my sincere thanks to Lauren Clarke and Abbey Persons for their help with data collection, as well as to all of the research participants that participated in this project. This project would not have been possible without them. Likewise, I am especially grateful to Jena Green and Beth Nisar for their considerable assistance in preparing this paper for publication.

Finally, I would like to acknowledge my family for their continued support.

TABLE OF CONTENTS

LIST OF FIGURES

Introduction

Clinicians have long employed perceptual evaluation as their primary tool during clinical assessment and diagnosis. Many speech symptoms indicate the nature and extent of neurological damage, an understanding of which is essential for accurate diagnosis. During assessment, clients are asked to perform various speech, non-speech, and speech-like tasks to reveal valuable information about the integrity of the speech mechanism. Clinicians can then form hypotheses about the underlying neuropathologies that result in disordered speech. This knowledge is needed in treatment planning and intervention. The degree of connection between what can be observed and what can be measured using a computer is one of the main considerations of the present study.

In research, speech may be analyzed perceptually, acoustically, or kinematically. Perceptual methods rely on listening skills for the assessment of speech. Acoustic analysis involves measuring specific characteristics of audio recordings, usually through specialized software and statistical analyses. Kinematic measures of articulatory movement may be obtained by attaching sensors to the articulators and recording their trajectories during speech (Dromey et al., 2018). While acoustic and kinematic measures are often used in research, perceptual analysis is most prevalent in clinical practice. Perceptual analysis is the primary means of data collection for clinicians during motor speech assessment. It is advantageous because it is efficient, inexpensive, and readily available in any clinical setting. It requires no special equipment because it is solely based on observation of the patient performing assessment tasks. Because one of the primary goals of communication is to be understood by listeners, it has the highest ecological validity (Wannberg et al., 2015). It also carries value for patients and their families.

However, Kent (1996) identified common limitations of perceptual evaluation. Perceptual judgments are subjective and are qualitative in nature, making them subject to human error. A primary concern is that there is no broad consensus on which perceptual dimensions should be rated and little agreement on how this should be done. Another complication is that perceptual features are not unidimensional, making it difficult to isolate specific characteristics for evaluation. Researchers have found variable intra- and inter-rater reliability in perceptual data (Wannberg et al., 2015). Sources of this variability include auditory illusions, where the brain misinterprets the auditory signal and hears something that is not there (Kent, 1996). One auditory illusion called the McGurk effect illustrates how visual input can confound a listener's perception. While listening to a recording of the sound /bʌ/, listeners viewed a speaker producing the sound /gʌ/. The conflicting visual and auditory information caused listeners to perceive the sound as /dʌ/; this highlights the unreliable nature of the human ear (Kent, 1996). As such, the validity and reliability of perceptual measures remain in question, and clinicians have expressed concerns over their efficacy in clinical practice (Chiu et al., 2021; Kent, Weismer, et al., 1999). While perceptual analysis is an established practice in the field, clinicians have sometimes turned to acoustic data to provide objective measures of speech.

It is difficult, however, to establish the connection between perceptual and acoustic data (Kent, 1996). A study by Chiu et al. (2021) that investigated these measures in individuals with Parkinson's disease suggested a modest and multifactorial relationship. Further, expert listeners often differ widely when associating perceptual features with a fixed set of acoustic metrics. At most, clinicians can observe general patterns in the data, but more specific conclusions are difficult to draw since several characteristics of disordered speech may have similar acoustic metrics but different etiologies. For example, reduced speed, range, and precision are common to both upper motor neuron damage and lower motor neuron damage (Karlsson et al., 2020; Kent, Duffy et al., 1999; Novotny et al., 2020). Additionally, some speech characteristics confound acoustic metrics. For example, there is no single acoustic measure that accurately represents articulatory precision or directly corresponds to intelligibility. It is therefore difficult to align acoustic metrics with perceptual ratings. Further research is necessary to better understand how to supplement perceptual measures with acoustic and kinematic data.

Acoustic analysis has been described as a valuable complement to traditional perceptual methods (Kent, Weismer, et al., 1999). Acoustic measures are valuable for establishing objective and consistent data on the characteristics of speech, particularly for research purposes. Acoustic data may include formant frequencies and their transitions, voice onset time, and vowel space area (Chiu et al., 2021). Through acoustic analysis, it is possible to identify deeper features of the speech signal that are not always evident to the human ear. Kent, Weismer, et al. (1999) discussed how acoustic analysis of diadochokinetic tasks can reveal characteristics specific to certain disorders that would not have been evident through auditory perception alone. For example, a perceptual study investigating syllable repetition in individuals with spastic dysarthria determined that syllables were produced at a normal rhythm and slower rate. Yet, an acoustic study of the same population identified abnormalities in rhythm (Kent, Weismer, et al., 1999). In another study, acoustic analysis determined that dysarthria related to Parkinson's disease is characterized by long syllables with incomplete closures, while Friedreich's ataxia most often presents short and slow syllables with complete closures (Kent, Weismer, et al., 1999). Acoustic metrics can provide a fuller picture of the disorder and assist clinicians in the diagnostic process.

While acoustic analysis can provide additional information about a patient's speech, it should be a support rather than a replacement for perceptual methods. Novotny et al. (2020)

found that acoustic algorithms could not match the sensitivity of human listeners. In addition, acoustic analysis is more time-consuming and labor intensive; therefore, it is not often used in clinical practice. With recent strides towards the development of automated acoustic analysis, and with this approach becoming more convenient and less expensive, acoustic analysis may become a more viable option for future use in the field (Kent, Weismer, et al., 1999; Rong, 2020). Kent, Weismer, et al. (1999) argued that a reliance on automated acoustic analysis could prevent clinicians from considering perceptual observations that may be clinically relevant. They advocate for using qualitative data to identify aspects that can be further analyzed using quantitative methods, suggesting a healthy balance should exist between acoustic and perceptual measures during clinical assessment and diagnosis. Combining qualitative and quantitative measures results in a more comprehensive clinical picture (Gadesmann & Miller, 2008). Because acoustic measures do not always closely correlate with perceptual descriptions of speech, it is not possible to quantify perceptual aspects of speech such as intelligibility or naturalness from acoustic metrics alone (Kent, 1996).

There are significant limitations to how much detail acoustic methods can reveal about the movements that underlie sound production. The concept of motor equivalence—that there are multiple combinations of articulatory movements that can be used to achieve the same acoustic goal—interferes with a clear understanding of speech movements through acoustic methods. Research conducted by Perkell et al. (1993) found evidence for the influence of motor equivalence in speech motor programming. When measuring the articulatory movements of typical speakers during production of the vowel /u/, Perkell et al. (1993) found significant variations in tongue-body raising and lip rounding, demonstrating that different movements can result in the same acoustic target. Motor equivalence thus leads to ambiguity in interpreting acoustic metrics because they do not directly measure the movements underlying the generation of sound. Because of this ambiguity in the acoustic signal, connecting acoustic measures to specific articulatory movements and possible underlying etiologies is challenging.

Kinematic data reveal the nature of speech movements and provide valuable insights into speech motor control. Through revealing the underlying movements of speech that are undetectable through perceptual or acoustic methods, kinematic measures allow researchers to draw conclusions about how the brain is controlling the speech mechanism. Wong et al. (2012) identified this as a significant advantage of kinematic data, as understanding the physiology of speech is helpful for detecting subclinical symptoms. They conducted a study to investigate the lingual kinematics of rapid syllable repetition tasks. Using acoustic methods, the researchers found similar rapid syllable repetition rates among healthy controls and individuals with motor speech disorders, but the kinematic data revealed significant differences in range, duration, and speed of movement. This illustrates the value of kinematic measures for identifying specific articulatory features that cannot be extracted from an acoustic signal but can lead clinicians to a greater understanding of an individual's motor speech system.

However, kinematic data present other challenges. Kinematic research relies on specialized equipment that is not widely available. In addition, kinematic data collection results in weaker ecological validity, as it requires the attachment of sensors to the articulators, which affects natural speech. After a brief period of habituation, the speaker may become accustomed to the sensors, but the speech remains altered to accommodate for their presence. Dromey et al. (2018) conducted a study to determine the amount of time necessary to allow for sufficient habituation. They found that a 10-minute period was necessary for speech to stabilize perceptually and acoustically after attachment of the sensors. Collecting kinematic data is time-consuming, expensive, and impractical for routine clinical use. Because of the resources required and the compromised ecological validity, kinematic analysis is conducted in research rather than clinical practice.

There are three main types of tasks that clinicians use during clinical assessment. Speech tasks, such as asking a patient to read a short passage or a sequence of increasingly complex words, hold the greatest ecological validity, as these tasks mirror typical speech behavior in a natural environment. Non-speech tasks involve activities that are unrelated to speech, such as applying pressure with the tongue or sustaining phonation. While they may reveal important information about neurological functioning, non-speech tasks carry the lowest ecological validity, as they are removed from the natural context of speech. Speechlike tasks approximate speech behaviors but do not qualify as true speech. They have higher ecological validity than non-speech tasks and aid clinicians in assessment and eventual diagnosis. For example, oral diadochokinetic (DDK) tasks involve rapid syllable repetition, which is used frequently in clinical assessment. Rong (2020) explained that oral DDK tasks are different from speech tasks in that they are less phonologically complex, require one's maximum speech rate, and do not hold communicative intent. DDKs may be employed as part of formal and informal assessments of motor speech (Gadesmann & Miller, 2008). Alternating motion rate (AMR) tasks are a subset of DDK, and involve repetition of the same syllable, such as /pʌpʌpʌ/, /tʌtʌtʌ/, or /kʌkʌkʌ/. The other type of DDK task is sequential motion rate (SMR), which involves repetition of a syllable sequence, such as /pʌtʌkʌ/.

Ben-David and Icht (2018) recognized the appeal of DDKs for use in clinical assessment, noting that these tasks are quick and simple to administer and noninvasive for the patient. Researchers have also used DDK tasks to examine the integrity of the speech mechanism, including identifying the maximum articulatory rate and gauging overall articulatory precision (Laskowski-La Morelle, 1994). There is no standard protocol among clinicians for administering DDK tasks, which presents an obstacle to clinical application (Laskowski-La Morelle, 1994). Typically, patients are provided with a model and instructed to produce the desired syllable sequence as quickly and steadily as they can (Kent et al., 2021). Clinicians observe the task and subjectively judge speech consistency and accuracy. DDKs are primarily measured in terms of articulation rate. Karlsson et al. (2020) determined that this is a clinically useful indicator of disordered speech. Clinicians also observe the steadiness and precision of articulatory movements. Clinicians can learn much simply from listening to patients perform DDK tasks. DDKs require the coordination of the sub-systems of respiration, phonation, resonance, and articulation (Allison et al., 2022). While performance on a DDK task does not necessarily predict intelligibility, it can shed light on an individual's overall articulatory function (Gadesmann & Miller, 2008). Yet, as mentioned previously, Gadesmann and Miller (2008) identified a lack of standard procedure and reliability measures for DDKs. The validity and reliability of perceptual judgments of rapid syllable repetition are important to determine, as rapid syllable repetition is a common task used in clinical assessment of motor speech disorders.

Articulation rate is a key component of clinical assessment. During DDK production, the patient is instructed to produce the target response at their maximum performance rate. As such, DDKs tax the motor system to a greater degree than speech tasks and can provide clinicians with important insights into neurological functioning (Rong, 2020). Rapid syllable repetition can therefore reveal information about the speed and stability of articulatory movement and aid clinicians during the assessment of articulatory precision (Kent, Duffy, et al., 1999; Laskowski-La Morelle, 1994; Wong et al., 2012). As DDKs are speech-like tasks that only approximate speech conditions, it is useful to examine articulation rate during speech as well. Therefore, the present study involves participants producing phonetically complex words at typical and rapid rates to allow measurement of speech movement parameters under habitual and challenging rate conditions. Redford (2014) found that it was possible to identify subtle changes in speech movements through analyzing speech rate in typically developing children. Smith et al. (1995) explored changes in articulation rate and the effect on speech in typical speakers. It was expected that repeating an utterance would result in greater stability due to the repeated activation of a single motor sequence. However, they found that as speakers moved away from their preferred speech rate into an unpracticed mode, the spatiotemporal index (a measure of variability across repetitions) increased, demonstrating less stability in articulatory movements. This provides compelling evidence that changes in speech rate are clinically significant.

In practice, clinicians rely on perceptual assessment, despite concerns about its accuracy and reliability. In this study we examined perceptual, acoustic, and kinematic data from rapid syllable repetition tasks and phonetically complex words to investigate the relationship between perceptual observations and acoustic and kinematic measures. We hoped this would shed light on how well perceptual ratings from listeners correspond to objective data. We anticipated natural variability in individual speakers' speed and precision of production. We hypothesized that ratings of precision would positively correlate with ratings of syllable steadiness. Finally, we predicted that speakers who were perceived as more precise would produce larger and faster articulator movements as measured from their acoustic and kinematic performance. A primary objective of the study was to learn whether perceptual, acoustic, and kinematic measures of rapid syllable production, as well as of complex word production, are congruent. We used correlation analyses to determine what, if any, relationships existed between these speech measures.

Method

All participants received a brief explanation of the study and signed a consent form approved by the Brigham Young University Institutional Review Board before participation.

Speaking Participants

Participants were 20 young adults between the ages of 18 and 29 (10 men ages 18–29 and 10 women ages 18–25) with no history of speech, language, or hearing disorders. All individuals were native speakers of Standard American English. Participants were recruited from the university community.

Listening Participants

Participants were 12 native English-speaking young adults (seven men and five women ages 22–29) from the university community with normal hearing.

Instrumentation

The speech samples used in this study were collected in the same session as recordings used in other research projects. All samples were recorded in a single-walled sound booth with the participant positioned 30 cm from a condenser microphone (AKG C2000B). A calibration recording at 50 cm was measured with a sound level meter to provide a point of reference for speech intensity in dB sound pressure level (SPL). An NDI Wave electromagnetic articulograph (Northern Digital Inc.) was used to collect articulatory kinematic data through the attachment of 3 mm sensor coils using cyanoacrylate adhesive to the following locations: two reference sensors on a set of eyeglasses that allowed for spatial orientation and corrected for head movement; two sensors at midline on the vermillion borders of the upper (UL) and lower lips (LL); two coils at midline at the tongue front (TF) and mid-tongue (TM); and one coil attached to the mandibular central incisors to measure jaw movement (J). Stomahesive (ConvaTec, Inc.) was used for the J sensor to protect the tooth enamel. The position of each sensor was tracked and recorded by a computer located outside the sound booth using the Wavefront software (Northern Digital Inc.). The kinematic data were sampled at 100 Hz and acoustic data at 22,050 Hz. For the listening component of the study, participants heard the speech stimuli via over-the-ear headphones (Sennheiser, HD 600) in a single-walled sound booth.

Speech Stimuli

The DDK tasks consisted of rapid syllable repetition for 5–10 seconds after a demonstration by the researcher. The participant was instructed: "Take a deep breath and repeat after me as quickly and steadily as you can. Keep going until I signal you to stop like this [thumbs up]." The tasks included the AMRs /pʌpʌpʌ/, /tʌtʌtʌ/, and /kʌkʌkʌ/, and the SMR /pʌtʌkʌ/. An SMR task was included because it was more complex and required greater articulatory coordination than the AMR task. The order of the stimuli was selected to progress from simpler to more complex.

We also chose to use challenging word-level stimuli to explore the relationship between perceptual, acoustic, and kinematic data in speech tasks for words produced at a habitual and a rapid rate. We selected three words from the *Apraxia Battery for Adults, Second Edition*: "disposability," "nationalistically," and "characteristically" (Dabul, 2000). These words were chosen because they are phonologically complex and were expected to elicit a range of articulatory performance. The words were presented five times in this order after the participant was given the following instructions: "You are going to see the words *disposability*, *nationalistically*, and *characteristically*. Please read aloud the words as you see them on the screen." Following this task, the participant was again presented with these same stimuli and instructed, "Please read aloud the words you see on the screen as fast as you can."

These words were chosen because they are phonologically complex, and it was anticipated that they would elicit a wide range in speech rate and precision across individuals.

The recorded speech samples were presented as stimuli for listeners to rate. Listeners rated each sample on speech precision, and the perceptual ratings were compared with the acoustic and kinematic data.

Speaking Procedure

After passing a hearing screening, participants had each of the sensors attached to their articulators and then were engaged in conversation for 5 minutes to allow for habituation to speaking with the sensors. The participants were then presented with the screener word and DDK stimulus sets. The screener word produced at a rapid rate always followed the screener word produced at a habitual rate to prevent the rapid repetition task from influencing typical performance. Ten participants produced the DDK repetitions first, and ten participants produced the screener words first.

Listening Procedure

In a custom MATLAB (Version2023a) application, participants were presented with example recordings from the speech samples selected by the researcher (see Figure 1). These recordings demonstrated speech at both ends of the spectrum for precision and steadiness. Listeners were allowed to familiarize themselves with the example recordings prior to beginning data collection. Participants were then presented with audio samples and asked to rate them for speech precision using a visual analog scale (VAS) in a custom MATLAB (Version 2023a) application. On the left end of the scale for both DDK and word samples was the word “casual” and on the right end of the scale was the word “crisp.” For the DDK samples only, a second VAS rating was completed for syllable steadiness, with the word “unsteady” at the left end and “steady” at the right end of the scale (see Figure 2). The speech samples were newly randomized for each rating, and the listener was allowed to rate each sample at their own pace and play back the recording as many times as desired, although they were not allowed to go back to previous samples once a rating had been saved.

Figure 1

MATLAB Application for Example Recordings

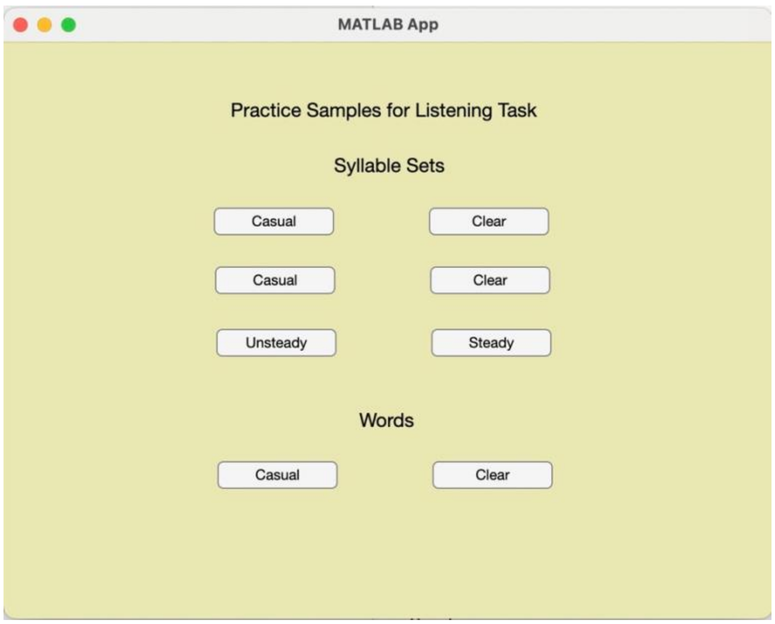
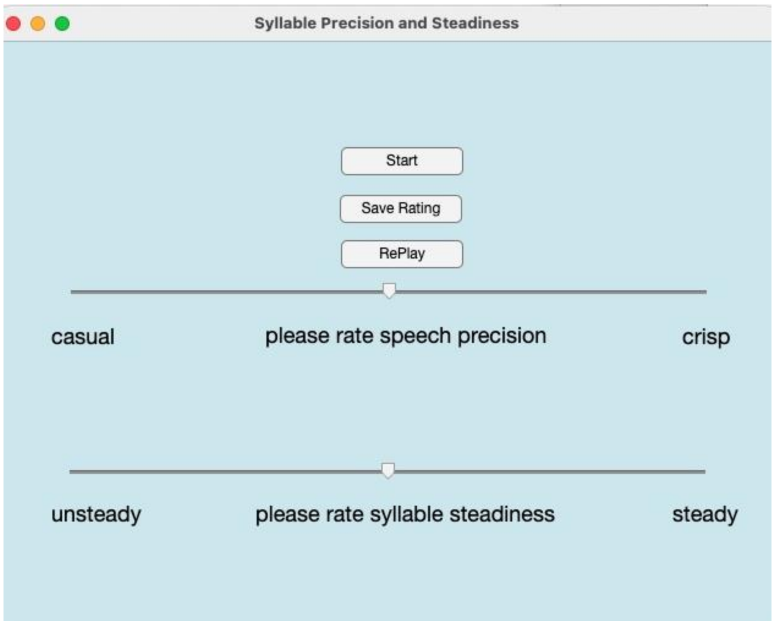


Figure 2

MATLAB VAS Application for Precision and Steadiness Ratings



Data Analysis

Perceptual

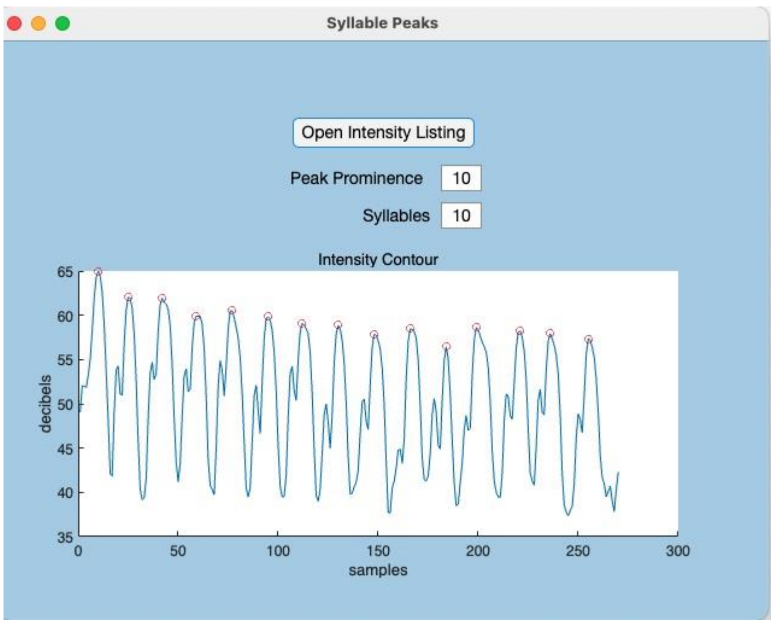
Raw data from the VAS app were saved in CSV files as a number between 0 and 100 for each audio sample. Intra-rater reliability was determined for each listener as the Pearson correlation between 20% of the samples that were randomly repeated and the ratings from the first time they were heard. The more reliable ratings (see results section below) were then tested for inter-rater agreement with an intraclass correlation coefficient. The average perceptual ratings served as the dependent variables for correlation with the acoustic and kinematic measures.

Acoustic

Praat software was used to extract an intensity contour from the first ten syllables in each DDK series. This intensity listing was imported into a custom MATLAB application which computed the duration of each syllable and the variability in the duration of syllables based on amplitude peaks for the vocalic element in each syllable (see Figure 3). This allowed a computation of the mean and standard deviation of the syllable rate and the coefficient of variation, reflecting the steadiness of the speech sample. The acoustic duration of each screener word was measured.

Figure 3

MATLAB Application for Syllable Metrics Computation

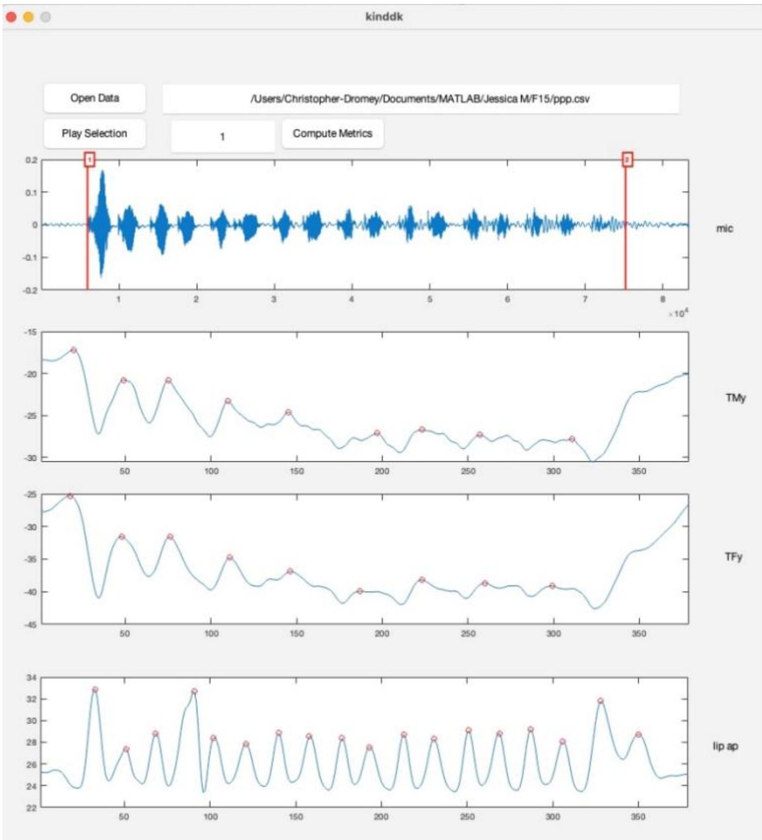


Kinematic

Kinematic data from the DDK tasks were analyzed by finding the displacement peaks for the tongue tip or tongue mid sensors for the /t/ and /k/ stop closures, respectively. Peaks in the lip aperture were automatically detected for the /p/ syllables (see Figure 4). Steadiness was determined by the consistency of the size of the movements and their peak velocities.

Figure 4

Lip Aperture Displacement Peaks for /pʌpʌpʌ/



Kinematic data for the screener words were derived through stroke metrics, a method outlined by Tasko and Westbury (2002) to separate speech kinematic records into measurable units. A MATLAB application calculated the Euclidean displacement of each sensor on a two-dimensional plane across the duration of the words. These sample-to-sample displacements were multiplied by the sample rate to create a speed history of each articulatory sensor, revealing the local speed minima (red circles in Figure 5). The software identified the articulatory strokes as occurring between successive speed minima in each articulator’s movement record. Two-dimensional x-y plots were used to calculate the hull area of each articulator for the duration of the word (see Figure 6), which reflects the area traversed by the articulator in the sagittal plane throughout the target utterance.

Figure 5

Lower Lip Speed Record for One Token of the Word 'Disposability'

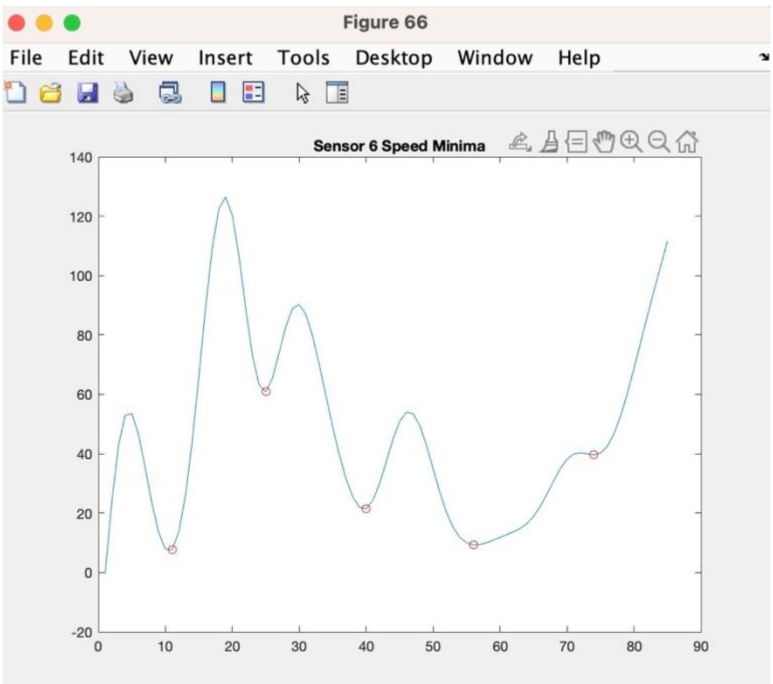
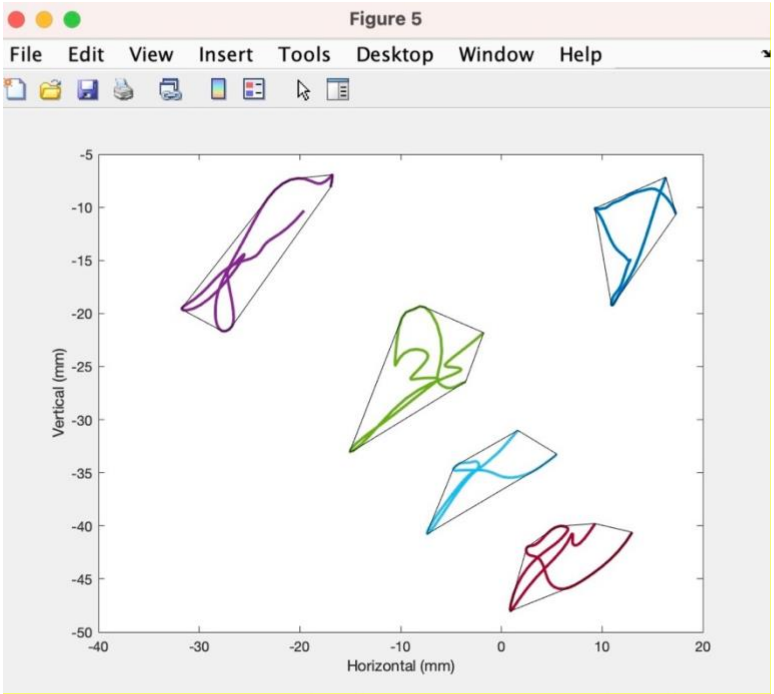


Figure 6

Hull Area for All Sensors for One Token of the Word 'Disposability'



Statistical Analysis

To quantify the association between the perceptual, acoustic, and kinematic measures, Pearson correlations were computed with SPSS (Version 29). Separate analyses were conducted for the DDK and screener word tasks.

Results

This study explored the relationship between perceptual, acoustic, and kinematic measures of speech. Kinematic variables included those for the tongue-mid (TM), tongue-front (TF), and lip aperture (LA). Significant correlations between perceptual, acoustic, and kinematic measures are reported below and displayed in the figures.

Reliability Testing

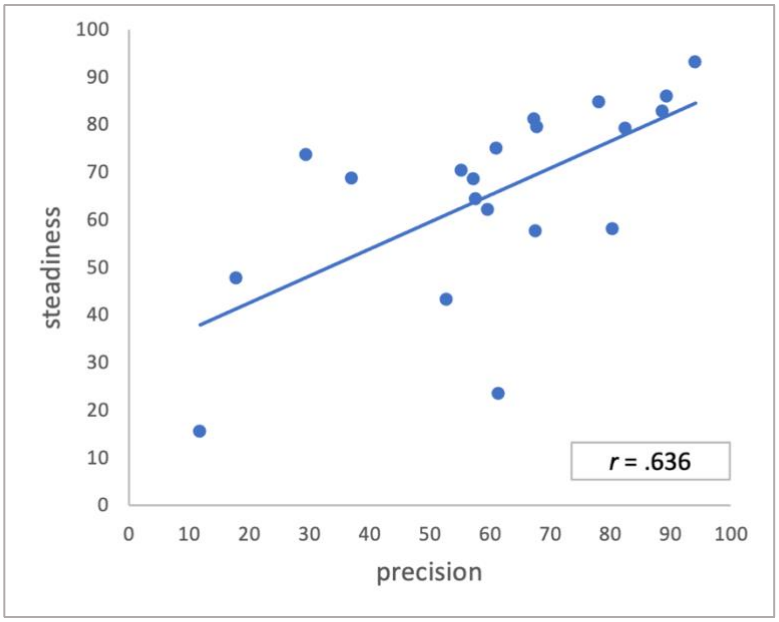
Perceptual ratings from randomly repeated recordings were compared with the original ratings to measure intra-rater reliability. Precision ratings from 10 of the 12 raters were included because their intra-rater reliability was above .50. The mean correlation of the reliable raters was .676 for precision and .723 for steadiness, demonstrating acceptable reliability. Fewer perceptual ratings of steadiness reached this level of reliability; thus, steadiness ratings from six raters were included in analysis. An analysis of rater agreement among the reliable raters revealed an average measures intraclass correlation coefficient of .881 for precision and .866 for steadiness, indicating good agreement among the raters.

Correlations Between Measures

For the /pʌpʌpʌ/ syllable task there was a significant correlation between perceptual ratings of precision and steadiness ($r = .636$, $p = .003$), as shown in Figure 7.

Figure 7

Precision Ratings vs. Steadiness Ratings for /pʌpʌpʌ/



There was a moderate negative correlation between perceptual ratings of steadiness and the standard deviation of the LA displacement for the /pʌpʌpʌ/ syllable task ($r = -.572$, $p = .008$), as shown in Figure 8, as well as the coefficient of variation of LA displacement ($r = -.657$, $p = .002$), as displayed in Figure 9.

Figure 8

Steadiness Ratings vs. Standard Deviation of Lip Aperture Displacement for /pʌpʌpʌ/

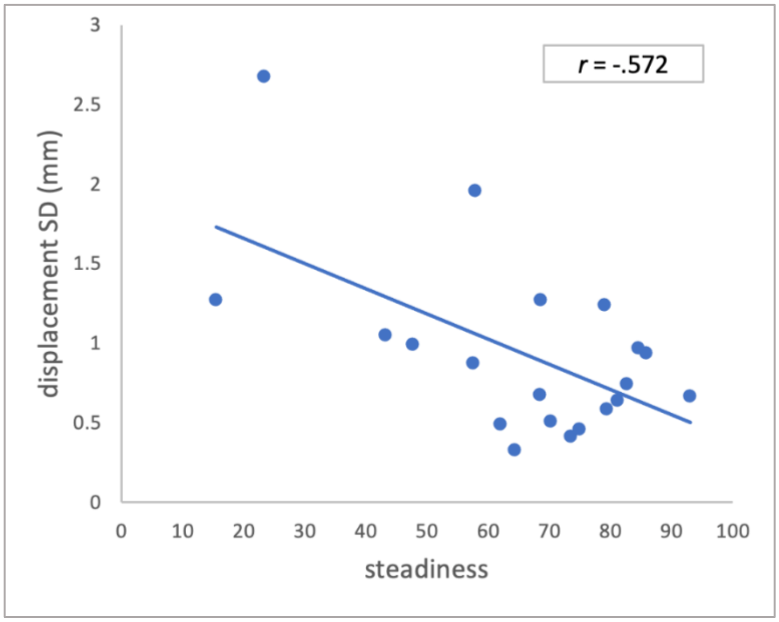
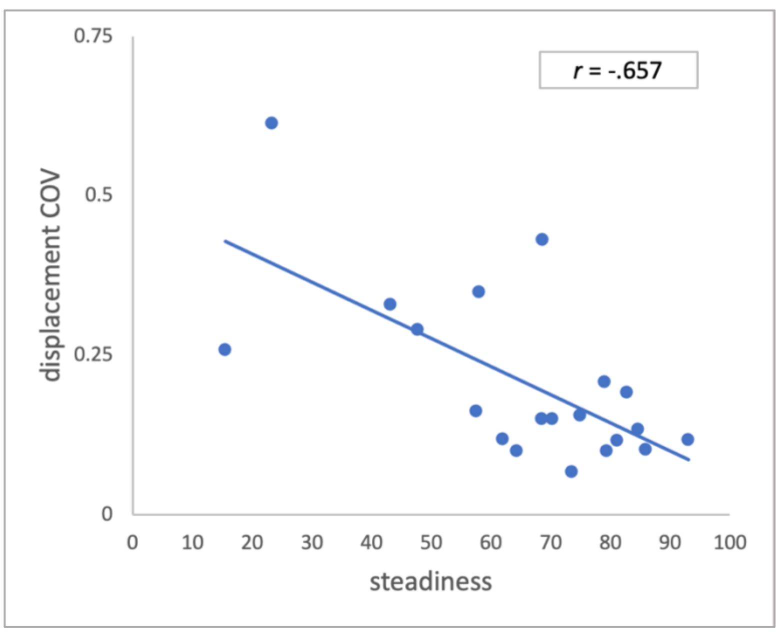


Figure 9

Steadiness Ratings vs. Coefficient of Variation of Lip Aperture Displacement for /pʌpʌpʌ/



As displayed in Figures 10 and 11, the mean LA velocity during the /pʌpʌpʌ/ syllable task demonstrated a moderate correlation with both precision (r = .559, p = .01) and steadiness (r = .577, p = .008) ratings.

Figure 10

Precision Ratings vs. Lip Aperture Velocity Mean for /pʌpʌpʌ/

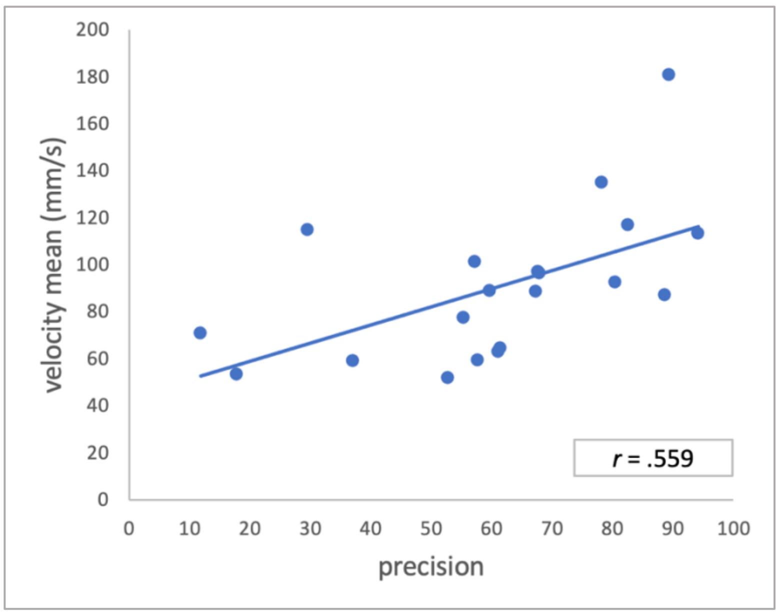
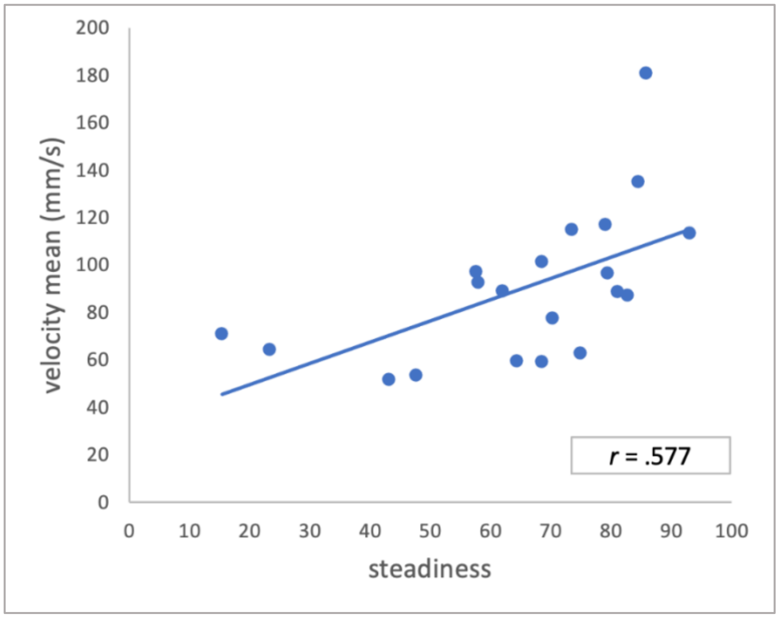


Figure 11

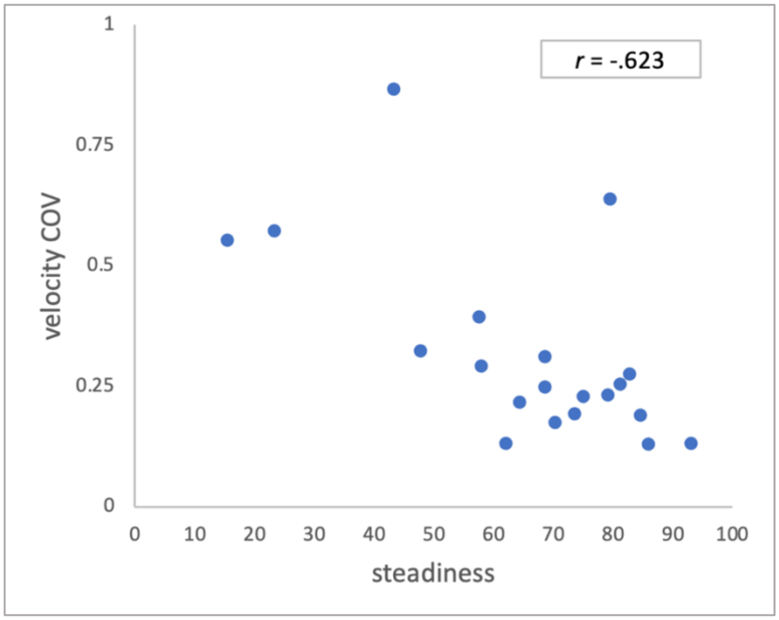
Steadiness Ratings vs. Lip Aperture Velocity Mean for /pʌpʌpʌ/



There was also a moderate correlation between ratings of steadiness and the coefficient of variation for LA velocity for /pʌpʌpʌ/ (r = -.623, p = .003), demonstrated in Figure 12.

Figure 12

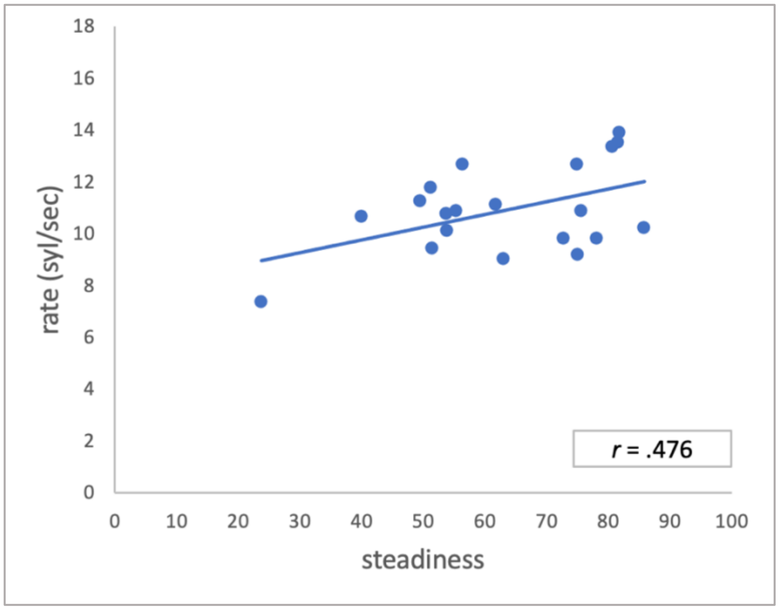
Steadiness Ratings vs. Coefficient of Variation of Lip Aperture Velocity for /pʌpʌpʌ/



For the /tʌtʌtʌ/ syllable task, a weak correlation was found between ratings of steadiness and acoustically measured rate (r = .476, p = .034), as shown in Figure 13.

Figure 13

Steadiness Ratings vs. Syllable Rate for /tʌtʌtʌ/



Ratings of steadiness also initially demonstrated moderate negative correlations with acoustically measured syllable duration standard deviation for /tʌtʌtʌ/ (r = -.622, p = .003) and coefficient of variation (r = -.597, p = .005). However, there were outliers in the data that affected the results. When these outliers were removed, the correlations were no longer found to be significant. These findings are displayed below in Figures 14–17.

Figure 14

Steadiness Ratings vs. Standard Deviation of Syllable Duration for /tʌtʌtʌ/

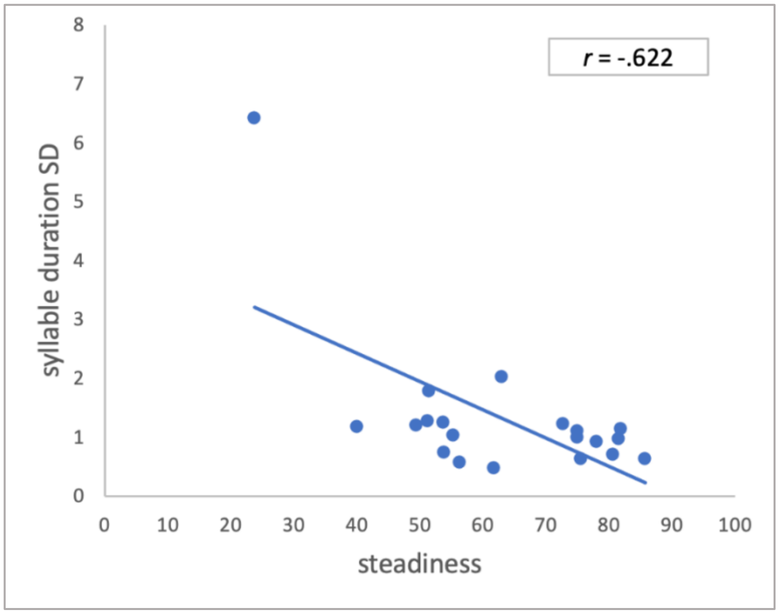


Figure 15

Steadiness Ratings vs. Standard Deviation of Syllable Duration for /tʌtʌtʌ/ Without Outlier

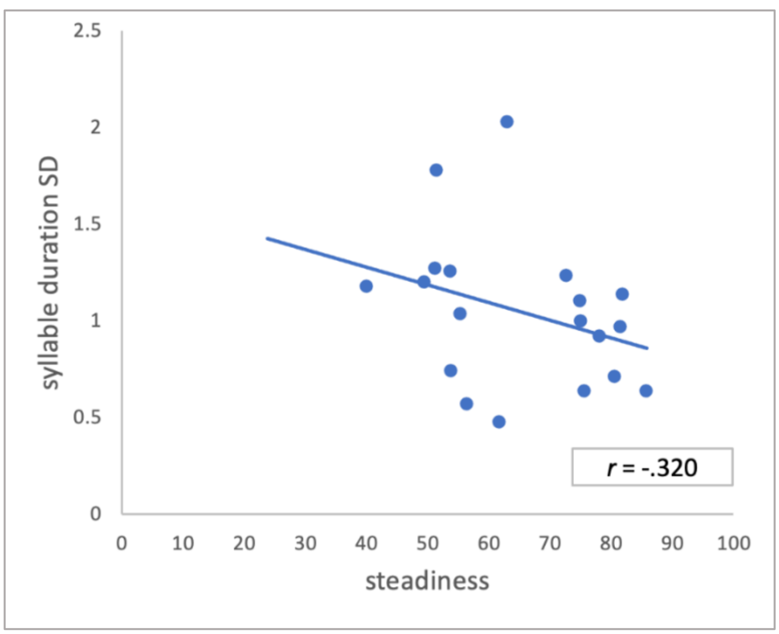


Figure 16

Steadiness Ratings vs. Coefficient of Variation of Syllable Duration for /tʌtʌtʌ/

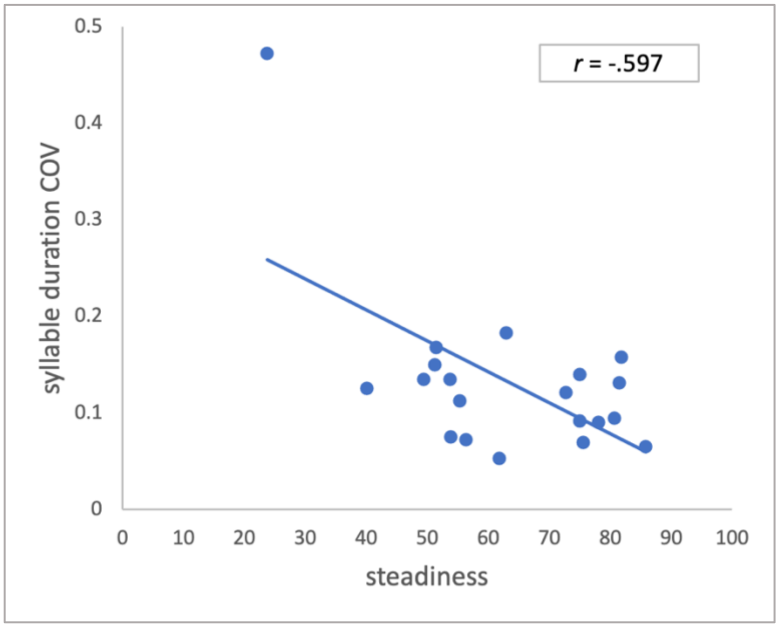
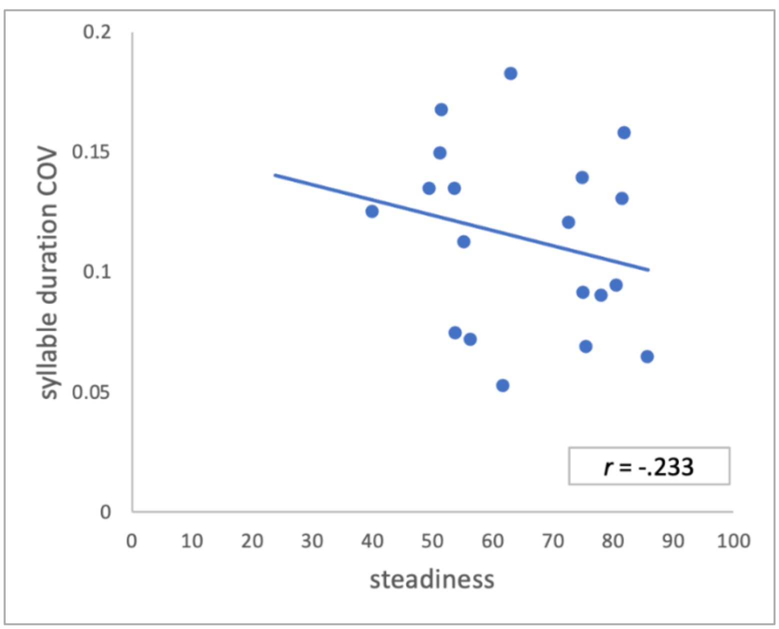


Figure 17

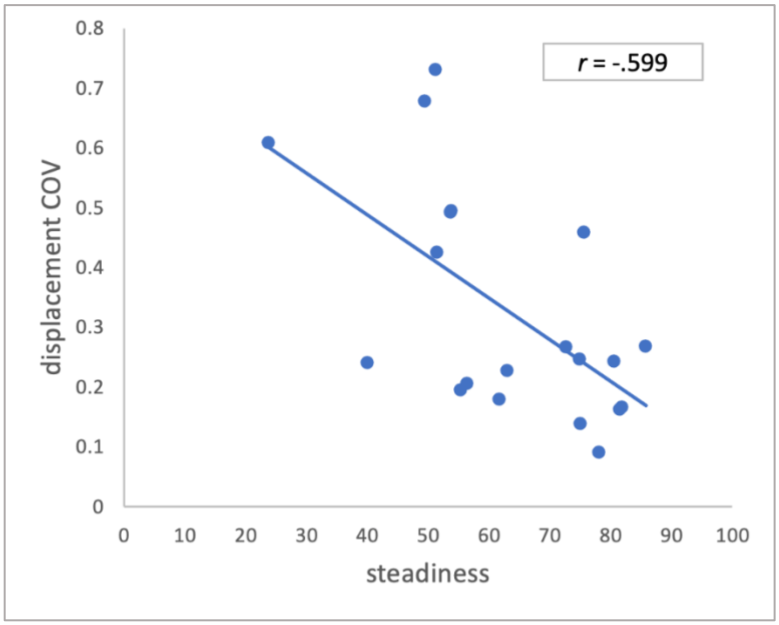
Steadiness Ratings vs. Coefficient of Variation of Syllable Duration for /tʌtʌtʌ/ Without Outlier



As shown in Figure 18, a moderate correlation between steadiness ratings of /tʌtʌtʌ/ and the coefficient of variation for TF displacement was also found (r = -.599, p = .005).

Figure 18

Steadiness Ratings vs. Coefficient of Variation of Tongue Front Displacement for /tʌtʌtʌ/



For the /kʌkʌkʌ/ syllable task there were moderate correlations between ratings of steadiness and the acoustic measures of syllable duration standard deviation (r = -.552, p = .012) and coefficient of variation (r = -.522, p = .018). However, once outliers were removed from the data the correlations were no longer significant. These results are displayed in Figures 19–22.

Figure 19

Steadiness Ratings vs. Standard Deviation of Syllable Duration for /kʌkʌkʌ/

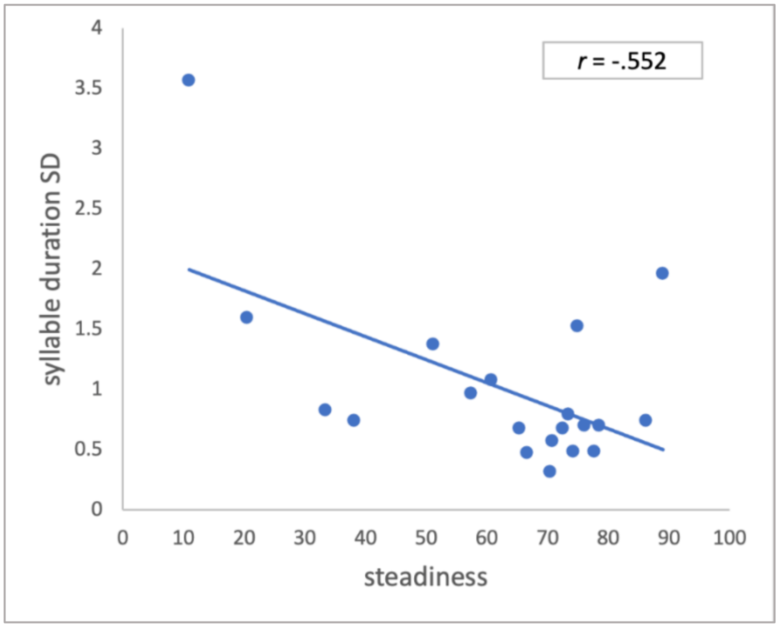


Figure 20

Steadiness Ratings vs. Standard Deviation of Syllable Duration for /kʌkʌkʌ/ Without Outlier

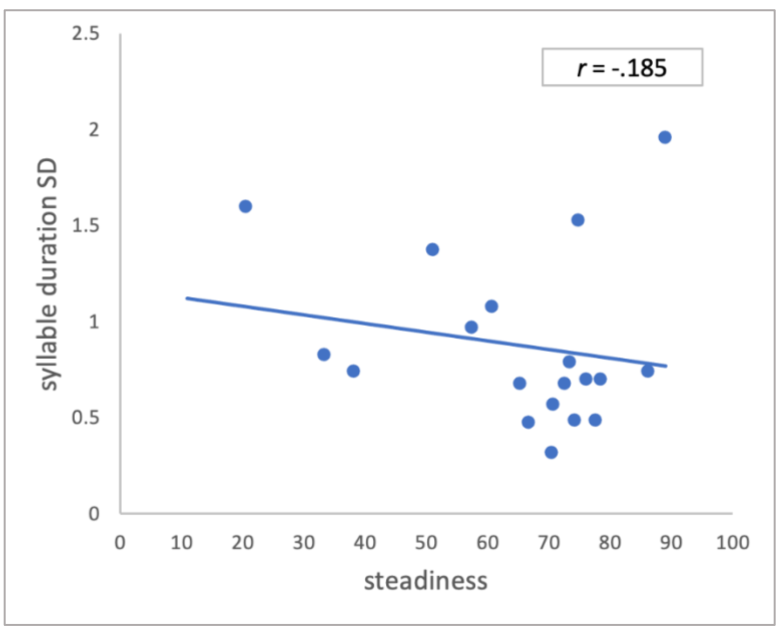


Figure 21

Steadiness Ratings vs. Coefficient of Variation of Syllable Duration for /kʌkʌkʌ/

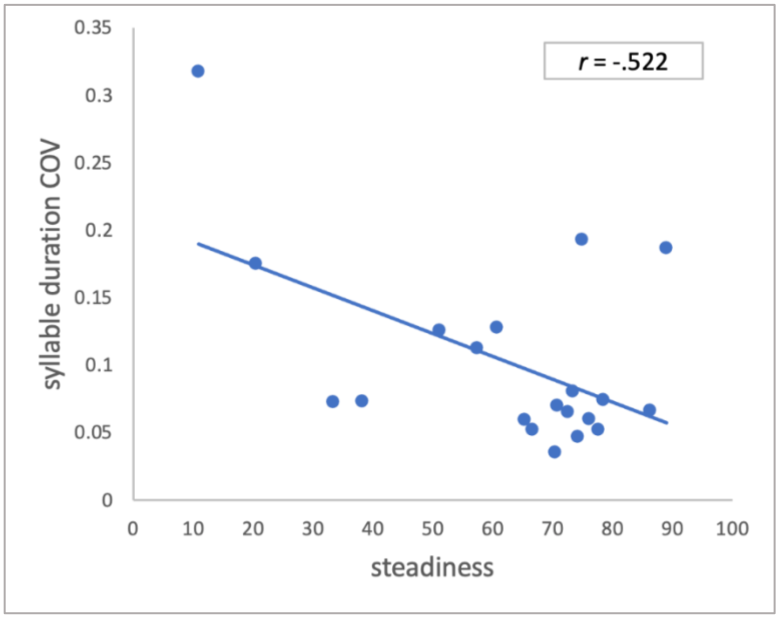
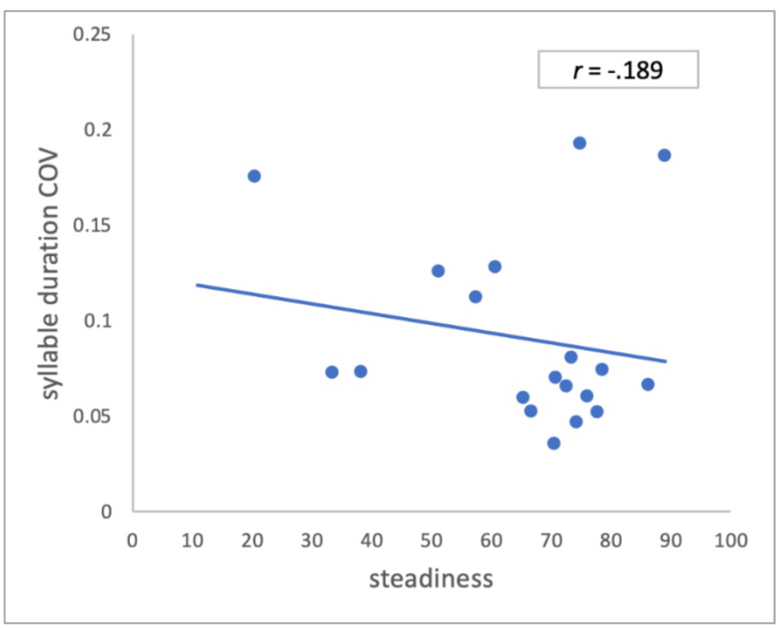


Figure 22

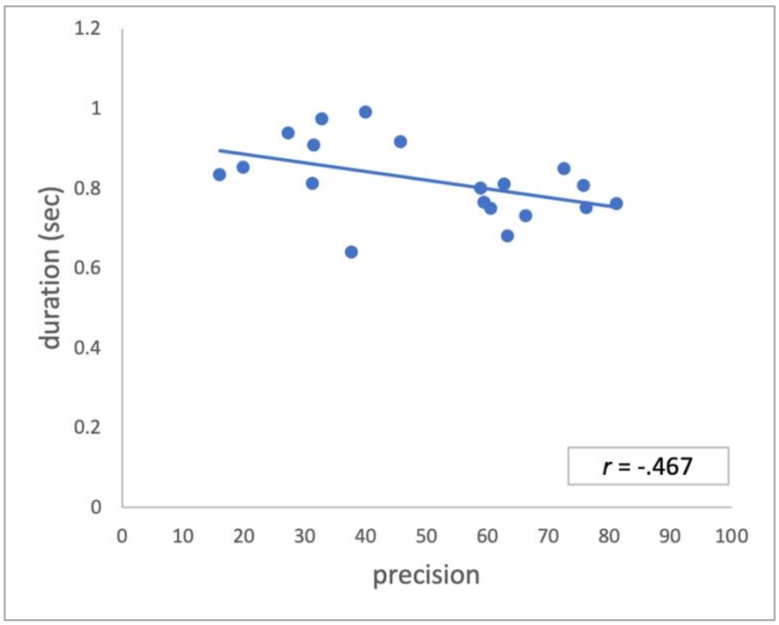
Steadiness Ratings vs. Coefficient of Variation of Syllable Duration for /kʌkʌkʌ/ Without Outlier



No significant correlations were found for the screener word at a regular rate. However, there was a weak negative correlation between perceptual ratings of precision of the screener word at a fast rate and acoustically measured duration ($r = -.467$, $p = .044$), as displayed in Figure 23.

Figure 23

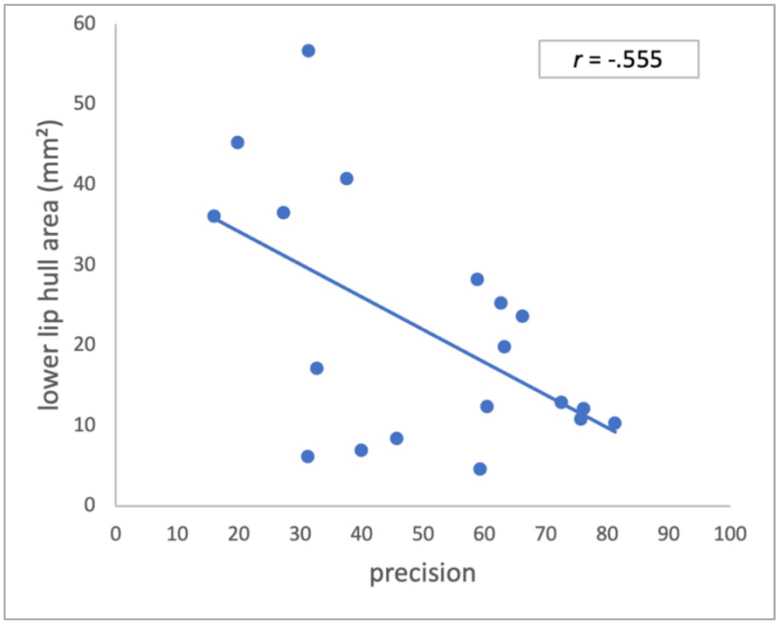
Precision Ratings vs. Word Duration for Screener Word (Fast)



There was also a moderate negative correlation between perceptual ratings of precision of the screener word at a fast rate and the kinematic measure of hull area for the lower lip ($r = -.555$, $p = .014$), as shown in Figure 24.

Figure 24

Precision Ratings vs. Lower Lip Hull Area for Screener Word (Fast)



Discussion

The purpose of this study was to investigate the relationship between perceptual, acoustic, and kinematic measures of speech. Analysis revealed several significant weak to moderate correlations between the three data types.

Reliability of Perceptual Ratings

For the DDK syllable tasks, listeners rated each recording for precision and steadiness because these perceptual measures have clinical utility in the differential diagnosis of dysarthria. However, because the screener words were not a repetitive stimulus item, they were only rated for precision.

Previous research on listener reliability has demonstrated mixed results. Gadesmann and Miller (2008) found low intra- and inter-rater reliability for DDK tasks, while Wannberg et al.

(2015) found that listeners demonstrated variable intra- and inter-rater reliability using assessment protocols for diagnosing motor speech disorders. Our findings were similar, as ratings for precision ranged in intra-rater reliability, with Pearson correlations from $r = .379$ to $r = .852$ and ratings for steadiness ranged from $r = -.438$ to $r = .942$. Some raters appeared to be inherently more reliable than others, despite being presented with the same protocol and speech samples. We found that raters were more reliable for precision than for steadiness, resulting in 10 raters (6 males and 4 females) having a reliability correlation for precision that was higher than $r = .50$ and 6 raters (4 males and 2 females) having a reliability correlation for steadiness that was higher than $r = .50$. However, the ratings for steadiness revealed more correlations with the objective data than the ratings for precision.

One possible explanation for the variability in reliability was that listeners were rating samples from typical speakers, which offered a narrow potential range of precision and steadiness. Correlations are typically weaker for datasets with a smaller range. For example, if researchers were to correlate weight and height among children, there would be a higher correlation with a sample of 5- to 15-year-old children compared to a sample of 5-year-old children. This could explain why the listeners in our study did not exhibit greater reliability, as the speech samples did not differ dramatically in steadiness and precision. Individuals with disordered speech would be expected to present a wider range of performance when compared to typical speakers in features such as steadiness, rate, and precision.

While the ratings for precision demonstrated a greater range overall, we found more correlations with objective measures for ratings of steadiness. These results could suggest that steadiness is an easier measure to rate perceptually, resulting in more correlations with acoustic and kinematic data. Perhaps the listeners felt that rating steadiness was more intuitive than distinguishing between "crisp" and "casual" speech, which are not common terms used by the general public to describe speech. Another possible explanation is that the speakers were more likely to make mistakes during production of DDKs, as they require a maximum rate and tax the motor system to a greater degree. Alternately, as Kent (1996) observed, mapping perceptual features directly to acoustic and kinematic data is challenging, and ratings for steadiness may simply have correlated better with the specific quantitative measures we chose for this study than ratings for precision.

Correlations Between Perceptual, Acoustic, and Kinematic Measures

We predicted that typical speakers would demonstrate natural variability in speed and precision for both the DDK tasks and screener words. This hypothesis was supported by the findings, as the speech samples displayed a measurable range in the acoustic and kinematic data and presented sufficient variability for listeners to rate for precision and steadiness along a continuum. However, listener reliability in distinguishing these differences varied greatly. This supports the principle of motor equivalence discussed by Perkell et al. (1993), in which the same acoustic target may be reached through different articulatory movements. The speakers in this study were given identical instructions, but kinematic data revealed that they may have used different articulatory strategies to produce the same target, resulting in natural variation that may be imperceptible to the human ear.

Perceptual ratings of precision and steadiness demonstrated a significant positive correlation with each other for the /pʌpʌpʌ/ syllable task, indicating these measures may be related. This supported our hypothesis that more precise speakers would also be more consistent. It is intuitive to expect that speakers who are perceived as more precise would also be perceived as producing steadier syllable repetitions. However, these results are limited to the AMR task

/pʌpʌpʌ/. This may be because /p/ is a labial sound and can be easier to distinguish perceptually than other sounds. Indeed, the greatest number of correlations between perceptual, acoustic, and kinematic data were found for the /pʌpʌpʌ/ syllable task. Also, bilabial closure is a simpler motor task than the tongue constriction required for alveolar or velar stops, which may have facilitated steadier production for the /p/ syllables.

Correlations between perceptual and acoustic data for DDK tasks were limited. For the /tʌtʌtʌ/ syllable task, ratings of steadiness positively correlated with acoustic rate, suggesting that as speakers increased their speed, they were perceived as being steadier. Karlsson et al. (2020) have shown that rate is a clinically useful measure in motor speech assessment, concluding that DDK rate was the best predictor for speech performance in individuals with Parkinson's disease. As we did not find significant correlations between rate and perceptual measures across DDK tasks, this may not extend to typical speakers. Nevertheless, our study supports a tentative connection between DDK rate and perceived steadiness. Smith et al. (1995) reported less consistent speech movements for slower speech during sentence production, which appears to be congruent with the current findings.

We initially found that ratings of steadiness demonstrated significant negative correlations with the acoustic measures of syllable duration standard deviation and coefficient of variation for both the /tʌtʌtʌ/ and /kʌkʌkʌ/ syllable tasks. This would have suggested that speakers who were perceived as more steady demonstrated less variability in syllable duration. However, outliers present in the data expanded the range dramatically, as shown above in the results section. When these outliers were removed, the results were no longer significant. Therefore, the only significant perceptual-acoustic correlation for DDK tasks was between ratings of steadiness and syllable rate for the /tʌtʌtʌ/ syllable task.

Likewise, there was only one significant correlation between perceptual and acoustic data for speech tasks. No significant findings were present for the screener word produced at a habitual rate. This could be due to the small sample size, as we only selected one word from each sample. Another possible explanation lies with the perceptual and kinematic measures we selected for this study. There are many different perceptual measures that could be compared and as Kent (1996) observed, there is little agreement as to which measures should be used clinically.

Ratings of the screener word produced at a faster rate showed an inverse relationship between precision ratings and word duration, indicating that faster speakers were perceived as more precise. Redford (2014) conducted a similar study that involved listeners rating typical speakers at their habitual rates. The data revealed that children with faster habitual rates were perceived as speaking more clearly, highlighting the connection between faster articulation rates and greater perceived speech accuracy. While this result was somewhat unexpected, this may be explained because a slower rate of speech is characteristic of some motor speech disorders. For example, Wong et al. (2012) found that speakers with dysarthria exhibited prolonged duration. Conversely, typical speakers are capable of producing intelligible speech more quickly and precisely. In our study, decreased duration for the screener word at a rapid rate reflected higher ratings of precision, suggesting a relationship between increased rate and speech precision.

We found preliminary evidence that perceptual ratings of steadiness may be correlated with acoustic DDK rate, and ratings of precision may be correlated with word duration in speech tasks in typical speakers. As there were few correlations between perceptual and acoustic measures across tasks, our findings support conclusions from previous researchers that only a modest relationship exists between perceptual and acoustic measures (Chiu et al., 2021). Further research is needed to explore the perceptual-acoustic relationship in the context of disordered populations.

The results for correlations between perceptual and kinematic data were more promising. Ratings of steadiness negatively correlated with the standard deviation and coefficient of variation of lip aperture displacement for the /pʌpʌpʌ/ syllable task. This means that speakers with less variability in lip aperture displacement measures were perceived as steadier. Likewise, steadiness ratings demonstrated a significant negative correlation with the coefficient of variation of tongue front displacement for the /tʌtʌtʌ/ syllable task. These correlations were found across DDK tasks and indicate that speakers with more consistent articulatory movements received higher steadiness ratings, which was a principal hypothesis of this study. Further, ratings of steadiness correlated with these kinematic measures but did not correlate with the comparable acoustic measures for these tasks. This indicates that kinematic data may reveal more or different information about the underlying speech mechanisms than acoustic data, supporting previous conclusions to that effect in the literature (Wong et al., 2012).

There was a significant positive correlation for ratings of precision and steadiness with the mean velocity of lip aperture for the /pʌpʌpʌ/ syllable task. This demonstrates that speakers with faster articulatory movements were rated higher for both precision and steadiness, which supports our hypothesis that speakers with faster articulatory movements would be perceived as more precise. Similarly, ratings of steadiness showed a significant negative correlation for coefficient of variation of lip aperture for /pʌpʌpʌ/ syllable task, demonstrating that speakers who exhibited less variability in velocity were perceived as steadier. Together with the kinematic data discussed above, these results suggest that less variability in displacement and velocity measures will translate to higher perceptual ratings of steadiness.

There were no significant correlations between perceptual and kinematic measures for the screener word at a habitual rate, perhaps for the same reasons mentioned above. Kinematic data from the screener word produced at a rapid rate revealed a significant negative correlation between ratings of precision and lower lip hull area. This correlation is counterintuitive, as it revealed that speakers with smaller movements of the lower lip were rated as more precise. In keeping with the literature, we would expect that larger movements would result in higher ratings of precision. Many interventions for individuals with dysarthria involve encouraging larger, over-articulated speech movements to improve intelligibility. The reasons for this unexpected finding are unclear.

Limitations and Directions for Future Research

This study's main limitation was using speech samples from typical speakers, which provided a limited range of precision and steadiness. This study could be profitably repeated with speakers with a variety of motor speech disorders to determine whether the observed linkages between perceptual and objective measures hold true for atypical speakers. The sample sizes of 20 speakers and 11 listeners were also relatively small. A similar study may be conducted with a larger sample size of both speakers and raters to reveal if this study's findings could be replicated and expanded given greater statistical power.

Another limitation was the presence of electromagnetic sensors attached to the articulators. Dromey et al. (2018) observed that attachment of sensors to the speech mechanism resulted in decreased perceptual ratings of speech precision, even after a period of time allowed for adaptation. Their study demonstrated that current methods of kinematic data collection affect natural speech and negatively impact ecological validity. As the aim of kinematic measures is to present an accurate representation of the movements involved in natural speech, this presents a methodological challenge. However, as there is no other suitable technology for collecting kinematic data, the attachment of sensors is a necessary compromise in obtaining the desired measures of articulatory motion.

Due to the limited scope of this study, we did not analyze data for the SMR syllable task. Similarly, we chose to analyze only one of the screener words. Future research could include a broader range of speech stimuli. Further, as DDKs are speechlike tasks rather than connected speech, the application of these findings to everyday speech remains limited. It would be beneficial to complete additional research using connected speech tasks to learn whether the relationships between objective and subjective speech measures hold true for more natural speech.

Conclusion

Speech analysis is integral to clinical assessment and diagnosis of motor speech disorders. Clinicians need reliable and valid measures of speech to make accurate diagnoses and inform intervention plans. Findings from this study are significant because they demonstrate that perceptual measures of speech may be correlated with objective data, providing support for the widespread clinical practice of relying on perceptual measures to assess and diagnosis motor speech disorders. This study identified a modest relationship between perceptual judgments and acoustic measures. However, there were significant correlations between perceptual and kinematic data, which highlights the importance of kinematic research to inform clinical assessment and diagnosis of motor speech disorders. For /pʌpʌpʌ/ and /tʌtʌtʌ/ syllable tasks, steadiness ratings were found to increase as standard deviation of displacement decreased, indicating more consistent movements translated to steadier speech movements as perceived by listeners. We also found significant correlations indicating that speakers with faster articulatory movements are perceived as more precise. It is hoped that future research will shed further light on the complex relationship between perceptual, acoustic, and kinematic data to aid clinicians in providing the best quality of care for clients.

References

Allison, K. M., Nip, I. S. B., & Rong, P. (2022). Use of automated kinematic diadochokinesis analysis to identify potential indicators of speech motor involvement in children with cerebral palsy. *American Journal of Speech-Language Pathology, 31*(6), 2835–2846. https://doi.org/10.1044/2022_AJSLP-21-00241

Ben-David, B. M., & Icht, M. (2018). The effect of practice and visual feedback on oral-diadochokinetic rates for younger and older adults. *Language and Speech, 61*(1), 113–134. https://doi.org/10.1177/0023830917708808

Chiu, Y.-F., Neel, A., & Loux, T. (2021). Exploring the acoustic perceptual relationship of speech in Parkinson's disease. *Journal of Speech, Language, and Hearing Research, 64*(5), 1560–1570. https://doi.org/10.1044/2021_JSLHR-20-00610

Dabul, B. L. (2000). *Apraxia Battery for Adults, Second Edition* (ABA-2). Pro-Ed.

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

Gadesmann, M., & Miller, N. (2008). Reliability of speech diadochokinetic test measurement. *International Journal of Language and Communication Disorders*, *43*(1), 41–54. https://doi.org/10.1080/13682820701234444

Karlsson, F., Schalling, E., Laakso, K., Johansson, K., & Hartelius, L. (2020). Assessment of speech impairment in patients with Parkinson's disease from acoustic quantifications of oral diadochokinetic sequences. *The Journal of the Acoustical Society of America, 147*(2), 839–851. https://doi.org/10.1121/10.0000581

Kent, R. D. (1996). Hearing and believing: Some limits to the auditory-perceptual assessment of speech and voice disorders. *American Journal of Speech-Language Pathology*, *5*(3), 7–23. https://doi.org/10.1044/1058-0360.0503.07

Kent, R. D., Duffy, J., Kent, J. F., Vorperian, H. K., & Thomas, J. E. (1999). Quantification of motor speech abilities in stroke: Time-energy analyses of syllable and word repetition. *Journal of Medical Speech-Language Pathology*, *7*(2), 83–90. https://mayoclinic.elsevierpure.com/en/publications/quantification-of-motor-speech-abilities-in-stroke-time-energy-an

Kent, R. D., Kim, Y., & Chen, L. (2021). Oral and laryngeal diadochokinesis across the life span: A scoping review of methods, reference data, and clinical applications. *Journal of Speech, Language, and Hearing Research, 65*(2), 574–623. https://doi.org/10.1044/2021_JSLHR-21-00396

Kent, R. D., Weismer, G., Kent, J. F., Vorperian, H. K., & Duffy, J. R. (1999). Acoustic studies of dysarthric speech: Methods, progress, and potential. *Journal of Communication Disorders*, *32*(3), 141–86. https://doi.org/10.1016/s0021-9924(99)00004-0

Laskowski-La Morelle, M. M. (1994). *Articulatory coordination adjustments as a function of increasing syllabic rate.* (Publication No. 421368) [Doctoral dissertation, Columbia University]. ProQuest Dissertations & Theses Global. https://www.proquest.com/comdisdome/docview/304123500/fulltextPDF/6A33627190024DA2PQ/1?accountid=4488

MathWorks. (2023). *MATLAB* (Version 2023a) [Computer software]. https://www.mathworks.com/products/new_products/release2023a.html

Northern Digital Inc. (2017). *Wave* [Electromagnetic articulograph system]. https://www.ndigital.com/

Novotny, M., Melechovsky, J., Rozenstoks, K., Tykalova, T., Kryze, P., Kanok, M., Klempir, J., & Rusz, J. (2020). Comparison of automated acoustic methods for oral diadochokinesis assessment in amyotrophic lateral sclerosis. *Journal of Speech, Language, and Hearing Research, 63*(10), 3453–3460. https://doi.org/10.1044/2020_JSLHR-20-00109

Perkell, J. S., Matthies, M. L., Svirsky, M. A., & Jordan, M. I. (1993). Trading relations between tongue-body raising and lip rounding in production of the vowel /u/: A pilot "motor equivalence" study. *Journal of the Acoustic Society of America*, *93*(5), 2948–2961. https://doi.org/10.1121/1.405814

Redford, M. A. (2014). The perceived clarity of children's speech varies as a function of their default articulation rate. *Journal of the Acoustical Society of America, 135*(5), 2952–2963. https://doi.org/10.1121/1.4869820

Rong, P. (2020). Automated acoustic analysis of oral diadochokinesis to assess bulbar motor involvement in amyotrophic lateral sclerosis. *Journal of Speech, Language, and Hearing Research, 63*(1), 59–73. https://doi.org/10.1044/2019_JSLHR-19-00178

Smith, A., Goffman, L., Zelaznik, H. N., Ying, G., & McGillem, C. (1995). Spatiotemporal stability and patterning of speech movement sequences. *Experimental Brain Research, 104*(3), 493–501. https://doi.org/10.1007/BF00231983

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)

Wannberg, P., Schalling, E., & Hartelius, L. (2015). Perceptual assessment of dysarthria: Comparison of a general and a detailed assessment protocol. *Logopedics Phoniatrics Vocology, 41*(4), 159–167. https://doi.org/10.3109/14015439.2015.1069889

Wong, M. N., Murdoch, B. E., & Whelan, B.-M. (2012). Lingual kinematics during rapid syllable repetition in Parkinson's disease. *International Journal of Language and Communication Disorders, 47*(5), 578–588. https://doi.org/10.1111/j.1460-6984.2012.00167.x

APPENDIX A

Annotated Bibliography

Allison, K. M., Nip, I. S. B., & Rong, P. (2022). Use of automated kinematic diadochokinesis analysis to identify potential indicators of speech motor involvement in children with cerebral palsy. *American Journal of Speech-Language Pathology, 31*(6), 2835–2846. https://doi.org/10.1044/2022_AJSLP-21-00241

Objective: This study examined articulatory movements to identify variables that may contribute to mild speech motor involvement in children with cerebral palsy. *Method:* Eight children with cerebral palsy and high intelligibility and eight typically developing children participated in an oral diadochokinesis (DDK) task by repeating the syllable "ba" for as long as possible in one breath. The participants' upper lip, lower lip, and jaw movements were recorded, and 23 kinematic measures were analyzed for diagnostic accuracy using receiver operating characteristic (ROC) analysis. *Results:* The authors found that five variables (duration of DDKs, number of cycles, and three measures relating to spatiotemporal data) demonstrated significant differences between the participant groups. The measures with the highest diagnostic accuracy (duration of DDK sequence and number of cycles) were shown to have 88% sensitivity and specificity. *Conclusion:* Automated methods for analyzing kinematic data from DDK tasks appear to differentiate between children with CP who demonstrate high intelligibility and children who are typically developing. *Relevance to the current work:* This study supports the use of DDK tasks in clinical assessment and diagnosis of motor speech disorders.

Ben-David, B. M., & Icht, M. (2018). The effect of practice and visual feedback on oral-diadochokinetic rates for younger and older adults. *Language and Speech, 61*(1), 113–134. https://doi.org/10.1177/0023830917708808

Objective: The authors examined the effect of visual feedback and motor practice on sequential motion rates (SMR) in younger adults and compared the results to the effects in older adults. *Method:* In a series of three experiments, younger and older adults were asked to repeat the SMR tasks /pʌtʌkʌ/ three times. The first experiment did not include visual feedback; the second experiment included visual feedback in the second of three rounds (with an auxiliary experiment including visual feedback in two of three rounds). The final experiment enhanced lighting and magnified the mirror to provide improved visual feedback during the second of two rounds. *Results:* Rates for younger adults were generally higher than for older adults. The first experiment found that performance in both groups was improved with the second practice but showed no change on the third. The second experiment demonstrated that visual feedback caused no change in older adults' performance and negated the practice round advantages for younger adults. The third experiment with improved visual feedback substantiated these results. *Conclusion:* SMRs were improved after a single practice round. Visual feedback was detrimental to younger adults' performance on SMR tasks and did not affect older adults' performance. *Relevance to the current work:* This study analyzed the rates of oral diadochokinetic tasks in typical adults.

Chiu, Y.-F., Neel, A., & Loux, T. (2021). Exploring the acoustic perceptual relationship of speech in Parkinson's disease. *Journal of Speech, Language, and Hearing Research, 64*(5), 1560–1570. https://doi.org/10.1044/2021_JSLHR-20-00610

Objective: The authors explored the relationship between acoustic and perceptual features of speech in individuals with Parkinson's disease (PD). *Method:* Speech samples from 13 individuals with PD and five typical speakers were rated by 20 listeners on four characteristics: ease of understanding, articulatory precision, voice quality, and prosodic adequacy. Ten acoustic measures were analyzed in relation to these perceptual measures using regression analyses. *Results:* Ratings of ease of understanding, articulatory precision, and voice quality for individuals with PD were lower than those for the control group. These perceptual features were found to be related to the acoustic measures of articulation rate and F2 transitions. *Conclusion:* The authors established a modest relationship between acoustic and perceptual characteristics of disordered speech in individuals with PD. *Relevance to the current work:* This study suggests that perceptual judgments may be based, in part, on underlying acoustic measures.

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: The researchers examined the amount of time needed for speakers to adapt to perturbation caused by an electromagnetic articulograph. *Method:* Twenty typical English speakers were instructed to read a set of stimuli prior to and at set intervals after the attachment of five sensors designed to measure articulatory movement. Twenty listeners participated in a perceptual evaluation and rated the speakers on speech precision using a visual analog scale. *Results:* Through acoustic and perceptual analysis, the authors found that the attachment of sensors resulted in decreased speech precision. However, precision after a period of 10 minutes remained consistent. *Conclusion:* This study suggests that a

10-minute period is sufficient for optimal speaker adaptation to sensors, allowing increased accuracy in data collection. *Relevance to the current work:* The current study also used the Wave system to measure articulatory movements; therefore, the conclusions of this study as to the amount of time necessary for a speaker to adapt to the presence of the sensors is highly relevant.

Gadesmann, M. & Miller, N. (2008). Reliability of speech diadochokinetic test measurement. *International Journal of Language and Communication Disorders*, *43*(1), 41–54. https://doi.org/10.1080/13682820701234444

Objective: The authors explored the inter- and intra-rater reliability of oral diadochokinetic (DDK) tasks and investigated whether rating performance is affected by clinical experience. *Method:* Twelve DDK speech samples from speakers with various neurological speech disorders were rated by ten clinicians and twelve listeners. Samples were rated on timing for repetitions and qualitative features, and the ratings for timing were compared to objective measurements from sound spectrograms. *Results:* The authors found that inter- and intra-rater reliability values were low, especially those for qualitative measurements, and did not reach acceptable levels for use in clinical diagnosis. Experience did not appear to affect rating accuracy. *Conclusion:* This research suggests that clinicians should be cautious with using perceptual DDK ratings for diagnostic decisions and highlights the need for further research on the impact of experience on ratings, as well as potential modifications that may improve rating accuracy. *Relevance to the current work:* This study focuses on assessing the reliability of DDK ratings for use in clinical assessment, which is the primary focus of the current work.

Karlsson, F., Schalling, E., Laakso, K., Johansson, K, & Hartelius, L. (2020). Assessment of speech impairment in patients with Parkinson's disease from acoustic quantifications of oral diadochokinetic sequences. *The Journal of the Acoustical Society of America, 147*(2), 839–851. https://doi.org/10.1121/10.0000581

Objective: The authors examined whether severity of speech impairment in passage readings in individuals with Parkinson's disease may be predicted by performance on the oral diadochokinesis (DDK) task. *Method:* Sixty-eight individuals with Parkinson's disease were recorded completing passage readings and a sequence of DDK tasks. The DDK speech samples were rated by four SLPs and analyzed acoustically, and the data were compared to overall performance on passage readings. *Results:* This study found that DDK sequences of /ka/ were the best predictor for speech performance. Production rate was an important measure used in determining the level of speech impairment. *Conclusion:* DDK performance may be used to predict performance on passage readings and overall level of speech impairment. *Relevance to the current work:* Articulation rate of DDKs is related to speech rate and can be used to predict severity of speech impairment.

Kent, R. D. (1996). Hearing and believing: Some limits to the auditory-perceptual assessment of speech and voice disorders. *American Journal of Speech-Language Pathology*, *5*(3), 7–23. https://doi.org/10.1044/1058-0360.0503.07

Purpose of the paper: This clinically focused paper discussed the concerns surrounding the validity and reliability of auditory-perceptual judgments and proposed ways to improve these methods. The paper outlined sources of variability, which may include auditory illusions, misperceptions of speech, phonemic false evaluation, and the McGurk effect. Limitations of sensitivity and reliability for auditory-perceptual judgments were also explored. The author observed that there is little agreement on the selection of perceptual features to rate for normal and disordered speakers. Inconsistent methodology, differences among rating scales, and biases influenced by speaker and listener characteristics further complicate the issue. The author proposed that procedures of measurement and methods of analysis be carefully considered to improve auditory-perceptual judgments during evaluation. *Relevance to the current work:* This paper identified limitations of auditory-perceptual judgments and suggested ways to improve the congruence between perceptual and acoustic metrics, which is a major focus of the present study comparing auditory-perceptual ratings to acoustic and kinematic data.

Kent, R. D., Duffy, J., Kent, J. F., Vorperian, H. K., & Thomas, J. E. (1999). Quantification of motor speech abilities in stroke: Time-energy analyses of syllable and word repetition. *Journal of Medical Speech-Language Pathology*, *7*(2), 83–90. https://mayoclinic.elsevierpure.com/en/publications/quantification-of-motor-speech-abilities-in-stroke-time-energy-an

Objective: This study collected data on rapid syllable repetition in individuals with stroke and dysarthria and compared the data to that of healthy controls. *Method:* Twenty-eight individuals with stroke and demonstrating dysarthric speech participated in rapid syllable repetition of the syllables /pʌ/, /tʌ/, /kʌ/, and /sʌ/. Quantitative measures, including syllable rate, syllable duration, and acoustic energy, were collected using acoustic analyses. *Results:* The participants with dysarthria exhibited slower syllable repetition compared to rates in healthy adults and demonstrated temporal irregularities that resulted in variable performance in syllable duration. Acoustic energy was also detected during the intersyllable gap and attributed to imprecise consonant articulation. *Conclusion:* The authors concluded that the slow performance of individuals with dysarthria on rapid syllable repetition tasks was due to intersyllable pauses and lengthened syllables. The study also suggested that quantification of acoustic energy present during rapid syllable repetition tasks has the potential to be a useful form of analysis for dysarthria resulting from stroke. *Relevance to the current work:* This study presents preliminary data on rapid syllable repetition in a clinical population.

Kent, R. D., Kim, Y., & Chen, L. (2021). Oral and laryngeal diadochokinesis across the life span: A scoping review of methods, reference data, and clinical applications. *Journal of Speech, Language, and Hearing Research, 65*, 574–623. https://doi.org/10.1044/2021_JSLHR-21-00396

Objective: The authors conducted a review of the current research on oral and laryngeal diadochokinesis (DDK) in children and adults with both typical development and clinical diagnoses. *Method:* The authors searched multiple databases and sources such as PubMed and Google Scholar with terms related to oral and laryngeal diadochokinesis, including DDK, syllable repetition rate, and alternating and sequential motion rate. *Results:* The authors found 360 publications related to the topic to include in the review. *Conclusion:* The use of DDK in clinical practice and research is confirmed by consulting the literature included in this review. *Relevance to the current work:* This scoping review covers the current research on oral and laryngeal diadochokinesis across the life span in various populations and supports the use of DDK tasks for the current work.

Kent, R. D., Weismer, G., Kent, J. F., Vorperian, H. K., & Duffy, J. R. (1999). Acoustic studies of dysarthric speech: Methods, progress, and potential. *Journal of Communication Disorders*, *32*(3), 141–86. https://doi.org/10.1016/s0021-9924(99)00004-0

Purpose of the paper: This paper provided a summary of available methods for the acoustic analysis of dysarthric speech. The categories of acoustic analysis and the validity and reliability of acoustic measures were briefly discussed. The authors outlined the equipment and resources needed for an acoustic study. Recommended acoustic measures for phonation, articulation, and resonance that could lead to a better understanding of acoustic-articulatory relationships were also included. The potential for creating acoustic typologies for the dysarthrias was examined. *Relevance to the current work:* The topic of this paper, acoustic analysis in clinical assessment, forms a fundamental part of the current study investigating the relationship between acoustic, kinematic, and perceptual means of examining speech.

Laskowski-La Morelle, M. M. (1994). *Articulatory coordination adjustments as a function of increasing syllabic rate.* (Publication No. 421368) [Doctoral dissertation, Columbia University]. ProQuest Dissertations & Theses Global. https://www.proquest.com/comdisdome/docview/304123500/fulltextPDF/6A33627190024DA2PQ/1?accountid=4488

Objective: The purpose of the study was to measure changes in articulation due to increased speech rate. *Method:* Twenty healthy adults, 10 men and 10 women, participated in a syllable repetition task at the rates of two, four, and eight syllables per second. Data on voice onset time, intraoral pressure, and vowel, syllable, and pressure duration were analyzed using quantitative measures. *Results:* Maximal articulatory rate was estimated to be eight syllables per second. Voice onset time could be reduced to increase rate for /bi/ but not for /pi/, and vowel duration experienced the greatest reduction and temporal compromise across all rates. The total pressure event and peak pressure rise time increased proportionally to repetition rate. *Conclusion:* This study provides preliminary data for syllabic segments during a maximum repetition rate task. *Relevance to the current work:* This work examined quantitative variables of articulatory activity during rapid syllable repetition in a typical population.

Novotny, M., Melechovsky, J., Rozenstoks, K., Tykalova, T., Kryze, P., Kanok, M., Klempir, J., & Rusz, J. (2020). Comparison of automated acoustic methods for oral diadochokinesis assessment in amyotrophic lateral sclerosis. *Journal of Speech, Language, and Hearing Research, 63*(10), 3453–3460. https://doi.org/10.1044/2020_JSLHR-20-00109

Objective: The authors compared available acoustic measures for analyzing speech samples of oral diadochokinetic tasks (DDKs) from individuals with amyotrophic lateral sclerosis (ALS). *Method:* Speech samples of a sequential motion rate (SMR) task from 18 individuals with ALS and 18 healthy individuals were analyzed using four algorithms based on varying signal processing approaches. *Results:* All algorithms detected irregularity and slow rate in DDK samples of individuals diagnosed with ALS with a threshold of 20 ms, but they were not as reliable as human annotators. Accuracy was higher for healthy individuals than for those with an ALS diagnosis. The DSP-multistep detector was the best acoustic method for detecting the presence of ALS. *Conclusion:* Automated acoustic assessment of DDKs may be an effective tool for monitoring disease progression in individuals with ALS. *Relevance to the current work:* Acoustic methods may be used reliably during clinical assessment to detect irregularity and slow rate of speech during DDK tasks.

Perkell, J. S., Matthies, M. L., Svirsky, M. A., & Jordan, M. I. (1993). Trading relations between tongue-body raising and lip rounding in production of the vowel /u/: A pilot "motor equivalence" study. *Journal of the Acoustic Society of America*, *93*(5), 2948–2961. https://doi.org/10.1121/1.405814

Objective: The authors conducted a kinematic study to collect evidence to support the hypothesis that different articulatory movements can result in the same acoustic goal. Method: The authors measured the articulatory movements of lip rounding and tongue raising with an electro-magnetic midsagittal articulometer (EMMA) during production of the vowel /u/. *Results:* Three of four participants displayed weak negative correlations of lip rounding and tongue raising, while the fourth subject demonstrated positive correlations. *Conclusion:* The results suggest that motor equivalence may be one strategy of speech motor programming that provides a compromise between economy of effort and speech clarity. *Relevance to the current work:* This study provided preliminary evidence for motor equivalence, which highlights the need for kinematic data to reveal the movements underlying speech acoustics.

Redford, M. A. (2014). The perceived clarity of children's speech varies as a function of their default articulation rate. *Journal of the Acoustical Society of America, 135*(5), 2952–2963. https://doi.org/10.1121/1.4869820

Objective: The author explored whether differences in children's default articulation rate suggest variation in articulatory timing control or some other factor unrelated to speech. *Method:* Six utterances from typically developing children were measured for articulation rate and segment duration. Fourteen adult listeners then rated the speech samples for clarity (enunciation) and related acoustic measures were taken. *Results:* Segmental duration accounted for differences in children's default articulation rates rather than suprasegmental changes. Listeners perceived utterances that were produced more quickly as clearer than utterances produced at slower rates. *Conclusion:* The findings were consistent with the motor skills hypothesis, indicating that greater articulatory timing control results in a higher default articulation rate. *Relevance to the current work:* Analysis of speech rate can provide valuable insight into speech movements.

Rong, P. (2020). Automated acoustic analysis of oral diadochokinesis to assess bulbar motor involvement in amyotrophic lateral sclerosis. *Journal of Speech, Language, and Hearing Research, 63*(1), 59–73. https://doi.org/10.1044/2019_JSLHR-19-00178

Objective: The purpose of this study was to validate a new method of automated acoustic analysis of an oral diadochokinesis (DDK) task, and to extract two temporal features, tongue movement jitter and alternating tongue movement rate, from this task to determine their ability to detect early bulbar motor involvement present in amyotrophic lateral sclerosis (ALS). *Method:* Sixteen individuals with ALS and eighteen controls participated in a DDK task and a task to determine speaking rate. Acoustic and kinematic data were collected and analyzed using automated acoustic analysis. *Results:* The automated acoustic analysis was validated, as the acoustic measures of cycle-to-cycle temporal variability (cTV) and syllable repetition rate (sylRate) were correlated with the kinematic measures of tongue movement jitter and alternating tongue movement rate. cTV demonstrated the best diagnostic accuracy (80% sensitivity and 94% specificity). The analysis also overcame the problems presented by continuous voicing. *Conclusion:*

This study provided an initial validation of a new method of automated acoustic DDK analysis and overcame continuous voicing between syllables, previously a major obstacle, suggesting its potential clinical utility for detection of bulbar motor involvement. *Relevance to the current work:* The implications for using automated acoustic analysis in clinical diagnosis with greater accuracy are relevant to the current work's focus on acoustic analysis compared to perceptual analysis when identifying motor speech disorders.

Smith, A., Goffman, L., Zelaznik, H. N., Ying, G., & McGillem, C. (1995). Spatiotemporal stability and patterning of speech movement sequences. *Experimental Brain Research, 104*(3), 493–501. https://doi.org/10.1007/BF00231983

Objective: This study examined the stability and pattern of speech movements and investigated the effect of changes in rate on articulatory movements. *Method:* Seven typical young adult speakers were recorded saying the phrase "Buy Bobby a puppy" at their habitual rate. They were then instructed to repeat the task at slow and fast rates. The data were analyzed to generate a new index for spatiotemporal stability, the relative timing of the peak velocity of the three opening movements, and pattern recognition techniques. *Results:* Distinct movement patterns were found within subjects for each of the three rates of speech. *Conclusion:* The findings suggest that speech rate is a global parameter and that a specific motor sequence is programmed for each new rate. *Relevance to the current work:* This study revealed important insights into how speech rate affects motor programming.

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: The authors examined a method of identifying movement "strokes" for reliable kinematic measurement during speech. *Method:* Speech samples of an oral reading task were measured kinematically and segmented into strokes, defined as the interval between two local minima of an articulator's movement. Articulator fleshpoints were determined from these kinematic measures. *Results:* Kinematic features helped distinguish speech strokes from tongue blade and jaw movements during alveolar fricatives. The authors found no connection between acoustic timing of alveolar fricatives and the kinematic measurements of the strokes. *Conclusion:* Stroke metrics can be taken for connected speech without requiring any external referents outside of articulatory movements. Strokes do not directly correlate to speech targets but allow for general conclusions to be drawn regarding speech organization. *Relevance to the current work:* Through stroke metrics, kinematic measurements of the articulators may be taken during speech to reveal information about speech movements.

Wannberg, P., Schalling, E., & Hartelius, L. (2015). Perceptual assessment of dysarthria: Comparison of a general and a detailed assessment protocol. *Logopedics Phoniatrics Vocology, 41*(4), 159–167. https://doi.org/10.3109/14015439.2015.1069889

Objective: The purpose of this study was to determine the reliability of a detailed assessment protocol compared to a general assessment protocol used in clinical diagnosis of motor speech disorders. *Method:* Five experienced clinicians rated speech samples of text reading from 20 individuals with varying types and severities of dysarthria using both assessments. The results of the two assessments were then compared. *Results:* The general assessment protocol was found to have higher intra- and inter-rater reliability, while only the detailed assessment protocol was able to identify more specific audible symptoms, such as imprecise consonants, monotony, and harsh voice. *Conclusion:* The authors concluded that the general assessment protocol was suitable for identifying key problem areas and severity of dysarthria, but the detailed assessment protocol was a clinically useful complement for determining audible symptoms and speech intelligibility. *Relevance to the current work:* This study examined the reliability of two perceptual assessments of motor speech disorders, which is the primary focus of the current work.

Wong, M. N., Murdoch, B. E., & Whelan, B.-M. (2012). Lingual kinematics during rapid syllable repetition in Parkinson's disease. *International Journal of Language and Communication Disorders, 47*(5), 578–588. https://doi.org/10.1111/j.1460-6984.2012.00167.x

Objective: The authors examined and compared the lingual kinematics of rapid syllable repetition in speakers with dysarthria related to Parkinson's disease, speakers with Parkinson's disease without dysarthria, and healthy individuals. *Method:* Participants were recorded performing rapid repetition of the syllables /tʌ/ and /kʌ/, and tongue-tip and tongue-back movements were measured through electromagnetic articulography. *Results:* The researchers found that individuals with Parkinson's disease and dysarthria demonstrated the greatest duration of tongue movement. Compared to the typical population, the individuals with Parkinson's disease had increased range and prolonged duration. However, all three groups had similar speech rates. *Conclusion:* This study

suggests that the prolonged duration of lingual movement present in hypokinetic dysarthria may be due to increased range rather than decreased range and speed.

Relevance to the current work: This study examined the lingual kinematics of individuals with motor speech disorders.

APPENDIX B

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.**

APPENDIX C

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: ____________