Formant Frequency Transitions in the Fluent Speech of Adults who do and do not Stutter: Testing the over-reliance on feedback hypothesis

Kaitlin Arnold
Western Michigan University, arnold_kaitlin@yahoo.com

Follow this and additional works at: https://scholarworks.wmich.edu/masters_theses

Part of the Communication Sciences and Disorders Commons

Recommended Citation
https://scholarworks.wmich.edu/masters_theses/590

This Masters Thesis-Open Access is brought to you for free and open access by the Graduate College at ScholarWorks at WMU. It has been accepted for inclusion in Master’s Theses by an authorized administrator of ScholarWorks at WMU. For more information, please contact maira.bundza@wmich.edu.
FORMANT FREQUENCY TRANSITIONS IN THE FLUENT SPEECH OF ADULTS WHO DO AND DO NOT STUTTER: TESTING THE OVER-RELIANCE ON FEEDBACK HYPOTHESIS

by

Kaitlin Arnold

A thesis submitted to the Graduate College in partial fulfillment of the requirements for the degree of Master of Arts Speech Pathology and Audiology Western Michigan University June 2015

Thesis Committee:

Stephen M. Tasko, Ph.D., Chair
Helen M. Sharp, Ph.D.
James Hillenbrand, Ph.D.
A recent theory proposed by Civier and colleagues (Civer et al. 2010; Civer et al. 2013) developed a model of stuttering that implicates a faulty feedforward control system. The hypothesis suggests that stuttering results from relying too heavily on sensory feedback to guide speech movements. An overreliance on sensory feedback would result in subtle anomalies in fluent speech (such as slowed articulatory transitions) as well as overt stuttering behaviors (such as sound repetitions). The present study tested this general hypothesis by comparing articulatory transition rates of adults who do and do not stutter across casual and fast speech rates.

Participants included 26 adults who stutter (AWS) and 28 normally fluent speakers (NFS) drawn from the Walter Reed-Western Michigan University Stuttering Database. Acoustic measures of articulatory transitions between the initial bilabial plosive /b/ and subsequent vowel /æ/ for the test utterance “a bad daba” (/ə bædæbə/). Acoustic data were analyzed for the following five measures 1) overall speech rate, 2) first and second formant transition durations, 3) first and second formant transition rates. For all measures, AWS tended to produce longer formant transition durations and slower transition rates compared to NFS, but the differences were not significant. The findings suggest that, there are some differences in formant values for AWS compared to NFS which are not significant. In general, findings from the current study do not provide strong support for the overreliance on feedback hypothesis.
Copyright by
Kaitlin Arnold
2015
ACKNOWLEDGEMENTS

I would like to extend my most gracious thanks to those who helped guide me throughout this thesis project. Your constant words of encouragement and seemingly endless knowledge has been paramount to this project’s completion and to my education. Dr. Tasko, your mentorship from undergrad to graduate school has been my most valued educational experience. I would not be here without Dr. Sharp, who has been a constant throughout my educational career at WMU. I also could not have gotten through this thesis nor my schooling without the comic relief and challenging education provided by Dr. Hillenbrand. The three of you, my thesis committee, have helped me endlessly and I simply cannot thank you enough. I also need to thank my family and friends for their continuous support and encouragement.

Kaitlin Arnold
# TABLE OF CONTENTS

**ACKNOWLEDGEMENTS** ......................................................................................................................................................... ii  
**LIST OF TABLES** ............................................................................................................................................................... v  
**LIST OF FIGURES** ............................................................................................................................................................... vi  
**CHAPTER**

I. **INTRODUCTION** ............................................................................................................................................................... 1  
   Characteristics of Stuttering .................................................................................................................................................. 1  
   Theoretical Perspectives of Stuttering ................................................................. 2  
   The Directions into Velocities of Articulators (DIVA) Model .................... 3  
   Feedforward vs Feedback Motor Control ......................................................... 4  
   Advantages of DIVA ......................................................................................... 6  
   Stuttering and the DIVA Model .................................................................... 7  

II. **THEORIES OF STUTTERING ETIOLOGY** .......................................................................................................................... 11  
   Summary of Literature and Study Hypotheses ............................................. 14  

III. **METHODS** ...................................................................................................................................................................... 16  
   Participants ........................................................................................................ 16  
   Stuttering Severity ........................................................................................... 17  
   Acoustic/Physiological Data Recording ......................................................... 17  
   Data Analyses ................................................................................................. 18
Table of Contents -- Continued

CHAPTER

Utterance Duration/Speech Rate.................................................................18

First and Second Formant Transitions.......................................................19

Statistical Analyses.................................................................................21

IV. RESULTS..............................................................................................22

V. DISCUSSION..........................................................................................30

Limitations and Future Directions.............................................................33

APPENDICES............................................................................................35

A. Human Subject Institutional Review Board Approval Letter..................34

B. Mixed-Effects Model Details for all Measures.......................................36

REFERENCES...........................................................................................41
LIST OF TABLES

1. Means and Standard Deviations for each measure based upon diagnostic category and speech condition..................................................................................................................23

2. Overall model values for Wald Chi-Square statistic which were assessed based on a $p$-value of 0.01 to control for an overall experiment-wide $p$-value of 0.05........................................................................................................................................24
LIST OF FIGURES

1. Key features of the DIVA model for typical speech production ......................... 3
2. Key features of the DIVA model for people who stutter, as hypothesized by Civier et al. (2010) .................................................................................................................. 7
3. Simulated speech output of the DIVA model ......................................................... 8
4. Sound pressure waveform and wideband spectrogram of the test utterance ‘a bad daba.’ ......................................................................................................................... 20
5. Average speech rates with Standard Error bars ..................................................... 25
6. Average values for F₁ transition duration with Standard Error bars ...................... 26
7. Average values for F₂ transition duration with Standard Error bars ...................... 27
8. Average values for F₁ transition rate with Standard Error bars ............................. 28
9. Average values for F₂ transition rate with Standard Error bars ............................. 29
CHAPTER I
INTRODUCTION

Characteristics of Stuttering

Stuttering is a complex, enigmatic disorder which has confounded those who have studied it throughout history. Van Riper (1984) defined stuttering as an abnormal disruption in the forward flow of speech. This disruption can manifest in what are commonly referred to as the core behaviors of stuttering; prolongations (the stretching out of a speech sound), repetitions (of speech sounds, syllables, whole words, or phrases), and blocks (the cessation of airflow and/or articulatory movement). In addition to these core stuttering behaviors, people who stutter may also exhibit many concomitant physical behaviors including; associated body movements, word or phrase interjections, speech fragments, vocal and respiratory anomalies, unusual eye movements, and cardiovascular phenomena such as changes in heart rate and blood pressure (Bloodstein & Ratner, 2008). People who stutter also exhibit a variety of emotional reactions to their stuttering. These can include fear or dread of speaking situations, anxiety, panic, embarrassment, shame, and resentment (Yairi & Seery, 2011).

The onset of stuttering typically occurs between two and three years of age, a time when there is a rapid expansion in the length and complexity of both speech and language (Yairi & Ambrose, 2005). Incidence and prevalence rates of stuttering are similar across cultures and languages, and there is growing evidence from a variety of sources that genetics play a factor in the etiology of stuttering (Bloodstein & Ratner, 2008). The prevalence of stuttering across the world is nearly 1% of the population and estimates of lifetime incidence rates range from 4% to 11% of the population (Yairi & Seery, 2011). This discrepancy between prevalence and lifetime
incidence highlights the fact that the majority of children who begin to stutter will recover with or without therapy. For the one percent who continue to stutter, it is typically a chronic, lifelong condition. Therefore, understanding the etiology and optimal management of stuttering remains a significant concern for both researchers and clinician alike.

Theoretical Perspectives of Stuttering

Throughout the years, a variety of theories regarding the underlying cause of stuttering have emerged. These theories have included environmental (Johnson et al., 1959), psychological (Bloodstein, 1995), linguistic (Postma & Kolk, 1993) and speech motor explanations (Zimmermann, 1980). A number of theoretical accounts of stuttering have emerged recently that implicate primary deficits in sensorimotor control (Smith 1999; van Lieshout et al. 2004; Alm, 2004; Loucks & deNil 2006; Max et al., 2004; Civier, Bullock, Max & Guenther, 2013). Support for this viewpoint comes from a number of sources. First, there is a growing literature demonstrating that children and adults who stutter exhibit speech motor patterns that are different from the non-stuttering population, even during fluent speech (e.g. Smith et al. 2010; McClean, Tasko & Runyan, 2004). Second, recent neuroimaging studies have identified functional and structural differences between children who do and do not stutter (Chang et al., 2008; Chang & Zhu, 2013). These differences involve reduced connectivity beneath the motor regions of the face and larynx and in the basal ganglia thalamo-cortical circuitry known to be associated with sequential motor control and sensorimotor learning. Third, significant advancements in contemporary theories/models of speech production allow greater opportunities for testing of specific hypotheses about stuttering. For example, in recent years there have been a handful of published studies which have attempted to provide a mechanistic account of
stuttering within the Directions into Velocities of Articulators (DIVA) model, a computational model of speech production (Cai et al. 2012; Civer et al., 2010; Civier et al., 2013).

The Directions into Velocities of Articulators (DIVA) Model

Briefly explained, DIVA is a neural network-based computational model that attempts to account for the acquisition and control of speech production. Some of the key features of DIVA can are shown in Figure 1.

**Figure 1.** Key features of the DIVA model for typical speech production. The bold arrows represent the higher reliance on the feedforward control system. Here, the feedforward reliance is 0.85 and the feedback reliance is 0.15. Adapted from Guenther et al. (2006).

Unlike competing models of speech perception and production (e.g. Saltzman & Munhall, 1989; Liberman et al. 1967), the DIVA model posits that speech sounds are represented
within an *acoustic* rather than an articulatory planning space. Speech sequences are planned as acoustic targets and a speech sound map (described below) is used to construct appropriate motor commands to the articulators based upon those targets. The motor commands are the result of two interacting control systems; feedback and feedforward control. The feedback control system (right side of Figure 1) serves to monitor key sensory systems\(^1\) and adjusts motor commands in response to changing environmental conditions. The feedforward control system (left side of Figure 1) relies on an internal representation of the periphery in the form of a speech sound map. The map is developed through sensory feedback during the lengthy period of speech sound acquisition. Sounds are heard, attempts are made to imitate those sounds, and an ongoing tuning of the relationship between articulatory positions and acoustic output occurs within the CNS. Over time, this tuning process results in a highly specific mapping of the acoustic-articulatory relationship. The speech sound map is utilized by the feedforward control system, which allows for the fast, accurate, and highly skilled movements required for speech production.

**Feedforward versus Feedback Motor Control**

Evidence for the impact of feedback manipulation in the stuttering population has existed for some time, specifically for that of Delayed Auditory Feedback (DAF). Sparks and colleagues (2002) explored the influence of DAF and speech rate on stuttering, variables which are known to enhance fluency in persons who stutter. They found differences based on stuttering severity in that subjects with severe stuttering showed improvements at both casual and fast rates when experiencing DAF.

\(^1\) While DIVA includes both auditory and somatosensory feedback, for sake of simplicity, only the auditory system is provided in Figure 1.
Within the DIVA model, the relative contribution of feedforward versus feedback control varies over the course of speech motor skill acquisition. Utilization of feedforward control is like having a map of where the articulators are to be in order to produce a target. Feedback control, on the other hand, requires the speaker to hear and feel where their articulators are and what they produce, which is then followed by adjustments to achieve the target. Whereas the shift from feedback to feedforward control occurs, in part, due to a major drawback of reliance on feedback control; sensory information takes too long to reach central neural structures. An overreliance on feedback control would limit the speed at which motor behaviors can occur if accuracy is to be maintained.

Fast and accurate movements are a prerequisite for fluent speech production. However, during the early stages of speech production, the young speaker lacks a sufficiently detailed speech sound map and thus must rely heavily on the feedback control system. Sensory information (auditory and somatosensory) via the feedback system is used to generate speech motor commands and help refine the map that links vocal tract positions to the resulting sound. As the speech system matures, the reliance on feedback decreases, while the feedforward system is enabled to use the speech sound map that has been developed, resulting in rapid, accurate, and flexible speech movements. Although typical speech production relies heavily on the feedforward system, the feedback system remains important for addressing environmental changes which cannot be accounted for by the feedforward mechanism alone.

DIVA employs the following weighting formula to determine the relative contribution of feedforward ($\alpha_{ff}$) and feedback ($\alpha_{fb}$) control in the overall motor command, $M(t)$.

$$M(t) = \alpha_{ff}M_{feedforward}(t) + \alpha_{fb}M_{feedback}(t)$$
The two weighting factors must sum to 1.0 and for typically developed speech are 0.85 and 0.15 for feedforward and feedback control respectively (Tourville et al., 2008). If the relationship is not developed in the expected way, a host of problems can arise within the speech production mechanism which will be explored subsequently.

Advantages of DIVA

DIVA has been a dominant model of speech production over the past two decades (Guenther, 1995; Golfinopoulos, Tourville, & Guenther, 2009) and there have been a number of reasons for its popularity. One reason is that the model generates both acoustic and articulatory kinematic output, allowing testable hypotheses about peripheral speech production processes (Guenther et al., 1999). Another reason is that the key components of DIVA were developed with reference to known functions of neural structures, making the model neurobiologically plausible, thus allowing for specific predictions about functional neuroimaging studies of speech production and perception (Golfinopolous, Tourville, & Guenther, 2009). Because of this, DIVA helps provide a linkage between neuroimaging studies and peripheral physiological and acoustic studies of speech. DIVA models speech motor development and provides important insights into the acquisition of speech motor control. Finally, DIVA is sufficiently explicit to allow for testing of a variety of hypotheses regarding speech disorders such as apraxia of speech, dysarthria, and stuttering (Terband, Maassen, Guenther, & Brumberg, 2009; Civier et al., 2010).
Stuttering and the DIVA Model

Recently, Civier and colleagues (2010; 2013) developed a model of stuttering that implicates a faulty feedforward control system. This faulty system involves diminished timing signals that prompt the transition from one syllable to the next and is described in detail in Civier et al. (2013). Due to the problems with feedforward control, the model predicts that people who stutter would adopt a strategy that relies much more heavily on the feedback control system (Figure 2).

![Figure 2](image-url)

Figure 2. Key features of the DIVA model for people who stutter, as hypothesized by Civier et al. (2010). The bold arrows represent the higher reliance on the feedback control system. Here, feedforward reliance is 0.25 and feedback reliance is 0.75.
Rather than using a typical weighting formula (0.85 feedforward, 0.15 feedback), the stuttering version of DIVA operates with the majority of the motor command deriving from the feedback control system (0.25 feedforward, 0.75 feedback). This shift in the control strategy predicts observable effects in both fluent and disfluent speech of people who stutter.

The left panel of Figure 3, which is drawn from Civier et al. (2010), shows the simulated speech output of the non-stuttering DIVA model outlined in Figure 1. The right panel shows the speech output for the stuttering version described in Figure 2. Both simulate the fluent production of the word /bid/.

Figure 3. Simulated speech output of the DIVA model using the non-stuttering version from Figure 1 ($\alpha_{ff} = 0.85$, $\alpha_{fb} = 0.15$) on the left, and the stuttering version from Figure 2 ($\alpha_{ff} = 0.25$, $\alpha_{fb} = 0.75$) on the right for the fluent production of the token /bid/. The dashed lines represent the acoustic target area, while the solid line is the actual production from the two versions of DIVA where the weighting formula was manipulated. The filled circles with straight lines between them are taken from Robb and Blomgren (1997), who used the fixed time points of 0 msec (onset- left circle), 30 msec following onset (middle circle), and 60 msec following onset (right circle). Note that in the non-stuttering version, the 30 msec circle falls after the transition is completed, indicating that some of the steady-state formant values were included in the analysis. This practice skewed results in the comparison of rates of formant transitions for stuttering and non-stuttering speakers (Civier et al., 2010).
In DIVA, the acoustic targets for vocalic sounds take the form of time-varying formant patterns. The dashed lines represent the acoustic target region for the formants and the solid lines indicate the actual formant trajectories for the productions. The acoustic target region is based upon the information contained in the speech sound map. Here it can be seen that when the system is biased toward feedforward control, the formants fall within or very close to the target region. When the bias is shifted to a higher reliance on the slower feedback control system, the formants fall outside the target region that spans the C-V boundary, where the formant transition rate is very high. This rapid change is where it can be seen that the formant transition rate is slower for the stuttering than the non-stuttering model. So while the production of the word is fluent, the acoustic patterns of the stuttering version of DIVA are distinct from that of the non-stuttering model.

The stuttering version of DIVA (Figure 2) also includes an additional monitoring subsystem which evaluates the magnitude of acoustic errors during sound production. This error is defined as the total deviation of $F_1$-$F_3$ values from target regions. If this acoustic error exceeds a threshold value, the Excessive Error Detector within the monitoring subsystem will trigger a repair causing a motor reset of the initiation of the production segment. This resetting would continue as long as the error level determined by the monitoring system exceeds a threshold value, and would be observed as a part-word repetition.

The proposed study will test the hypothesis presented by Civier et al. (2010) by examining formant transition patterns in the fluent speech of stuttering and non-stuttering speakers. It is hypothesized that an overreliance on the feedback control system will appear as a slower rate of formant transitions when compared to normally fluent speakers. If this is the case,
clinical work in the area of stuttering, which takes advantage of rate modification techniques for fluency enhancing strategies would be further supported (Blomgren and Goberman, 2008).
CHAPTER II

THEORIES OF STUTTERING ETIOLOGY

There are a number of previously published studies that have attempted to examine the relationship between stuttering and formant transition patterns. Robb and Blomgren (1997) hypothesized that stuttering is associated with difficulties transitioning from sound to sound, and that this difficulty would result in differences in the articulatory and acoustic transitions at consonant-vowel boundaries. The authors examined the changes in second formant (F2) values during fluent production of CVt tokens in five AWS and five normally fluent speakers (NFS). F2 values were extracted at vowel onset and at fixed points of 30 and 60 msecs following vowel onset. Each F2 transition was represented as a slope. The AWS exhibited a trend to have larger slope values than the NFS for many of the test conditions. However, a great deal of variation was observed across vowel and consonant context making general conclusions difficult in these small samples. Figure 3, drawn from Civier et al. (2010) plots the Robb and Blomgren data for the test word /bid/. The symbols in each plot represent the average F2 value at each of the three time points (i.e. onset, 30 msec, and 60 msec). For the normally fluent subject data, it appears that at 60 msec post onset, the F2 transition is complete. For the data of subjects who stutter, at 60 msec post onset, the F2 continues to increase, indicating that the F2 transition is not yet complete. This example demonstrates how the data of subjects who stutter could exhibit a greater slope value than the non-stuttering data, even though the F2 transition for the stuttering data appears longer and more gradual. It is notable that the F2 transitions for the non-stuttering and stuttering groups roughly matches the simulated F2 transitions produced by the stuttering DIVA.
Blomgren, Robb, and Chen (1998) evaluated vowel space using the same subjects from the previously mentioned research. They tested the hypothesis that AWS exhibit a reduced vowel space when compared to normally fluent adults. The authors examined F\textsubscript{1} and F\textsubscript{2} values during the steady-state region of three corner vowels (/i, u, \alpha/) in a CV\textsubscript{t} context in a group of 15 adult males which included five untreated (within the last five years) and five treated AWS, along with five NFS. Only fluent productions were included in the analyses. Measures included 1) estimates of F\textsubscript{1}-F\textsubscript{2} spacing, 2) the area of the vowel triangle in the F\textsubscript{1}-F\textsubscript{2} space and, 3) the Euclidean distance from the centroid of the vowel triangle to the vertex associated with each vowel. Results indicated that, across selected measures, there was a trend for the untreated AWS to have a reduced formant space as compared with controls. Additionally, AWS also exhibited longer vowel durations when compared to the control group. However, results were quite variable across different vowels, utterance samples, and group sizes were small, thus making generalization difficult.

In an effort to determine if stuttering is associated with abnormal anticipatory coarticulation, Sussman, Byrd, and Guitar (2010) derived F\textsubscript{2} locus equations for fluent and non-fluent productions of stop+vowel stimuli in a group of eight AWS. Only five subjects were used in the analysis because three of them had too few disfluent productions. No normally fluent controls were used; instead the authors relied on previously published data. F\textsubscript{2} frequencies were measured at the onset and the visually determined midpoint of each vowel. The main measures of interest for this study were the coefficients (i.e. slopes and y-intercepts) of the locus equations, which are regression equations that fit the relationship between F\textsubscript{2} onset and F\textsubscript{2} midpoint. The results of the analyses showed that the stuttered productions had a considerably higher standard error of estimate (a numerical representation of the distribution of the data points around the
regression line), when compared to the fluent productions. However, the basic form of the locus equations did not clearly distinguish the stuttering group from previously published work on NFS. It should be noted that, as is the convention for the development of locus equations, the duration between the F2 onset and F2 midpoint was not reported, making it impossible to determine if there was evidence of a reduced rate of F2 transition.

An analysis of acoustic data for fluent and disfluent speech for F2 transitions in 13 children who stutter (CWS) was conducted by Yaruss and Conture (1993). The children were divided into two groups based on their likelihood for persistence (based on Stuttering Prediction Index scores). Five acoustic measures were made; duration of F2 transition, onset and offset frequencies of F2 transitions, extent of F2 transition, and rate of frequency change in F2 transition. The results indicated that formant transitions differ between fluent and stuttered productions both within and between groups (low-risk and high-risk for persistence of stuttering). However, since the study did not include a control group, it is not possible to determine if the F2 transition measures of the fluent productions of either group of CWS were different from non-stuttering peers.

In an effort to evaluate formant transition duration and rate, Zebrowski, Conture, and Cudahy (1985) compared the speech of 11 young CWS with 11 normally fluent counterparts. A large number of acoustic measures were made from subjects’ fluent speech including consonant-vowel F2 transition duration and rate. Results failed to show any significant differences between the two groups. The authors attempted to explain the lack of significant findings by identifying some inherent challenges associated with their study including the difficulty of obtaining reliable acoustic data for children as well as the small subject sample size.
The second formant transitions of 14 children who stutter and 14 fluent, age matched peers were examined by Chang, Ohde, and Conture (2002). Speech samples were words containing a variety of consonant-vowel transitions. The authors generated locus equations and determined second formant transition rates for the samples. Results showed no group differences in the locus equation analysis and failed to show any evidence that formant transition rates were different for CWS and normally fluent children. The authors did find that differences in formant transition rates based on place of articulation were not as marked for the CWS as compared to the normally fluent children, which the authors interpreted as evidence for a less refined speech motor organization in CWS.

Summary of Literature and Study Hypotheses

Overall, studies that have examined formant transitions in the fluent speech of people who stutter have produced mixed results. In the case of studies that employed the traditional locus equation method, the extent of formant transitions were measured but the durations over which the transitions occurred were not (Sussman et al., 2010; Chang et al., 2002). Although Robb and Blomgren (1997) included information regarding timing of formant transitions, their use of a fixed time-point criterion likely misrepresented the actual rates of formant transitions. Their results showed that formant transition rates of AWS were actually faster than NFS; however, this could be the result of the fixed times points being located beyond the completion of the transition. As plotted in Figure 3, these results imply a slower rate of transition for normally fluent speakers when it is actually the opposite; the transitions were too fast to be accurately captured using the fixed time-point method.
The proposed study will directly evaluate the hypothesis of Civier et al. (2010) by examining both F1 and F2 transitions in adults who do and do not stutter at casual and fast speaking rates. Specifically, transition duration and rate of both formants will be analyzed. Both formants will be evaluated because both F1 and F2 contribute to the error estimates in the stuttering DIVA. Speaking rate, which has not been controlled in previous studies, will be manipulated because Civier’s model predicts that people who stutter would exhibit a progressively more difficult time producing normal formant transition patterns as speaking rate increases. This is because the hypothesized problems associated with an over-reliance on feedback control would be most evident when speech production must be fast. It is predicted that compared to normally fluent speakers, AWS

1. Will exhibit significantly slower formant transition durations and rates when speaking at a fast speaking rate.

2. May exhibit significantly slower formant transition durations and rates when speaking at a casual rate.
CHAPTER III

METHODS

Participants

Participants were drawn from the Walter Reed Army Medical Center-Western Michigan University Stuttering Database. This database includes a range of clinical-behavioral and physiological data obtained on 43 adults who stutter (AWS) and 43 normally fluent adult speakers (NFS). Participants were largely reserve and active duty members of the United States Armed Services. All AWS reported that they had stuttered since childhood and were seeking stuttering treatment through the Walter Reed Stuttering Treatment Program. All NFS reported no history of speech or hearing difficulties and were briefly screened by a certified speech-language pathologist prior to inclusion in the study.

Not all records from the database could be used in the current study. For some participants, the number of usable replicates of the speech task of interest was not sufficient for analysis. A replicate was considered unusable if it was disfluent, misarticulated, or recording quality was not adequate to allow for formant extraction in the speech task of interest. Twenty-six AWS and 28 NFS, matched for age, ethnicity, and race were used in the analysis. Data for all subjects included F$_1$ transition duration and rate for both speech conditions. In some instances, the F$_2$ data were eliminated due to difficulty accurately tracking formant values.
Stuttering Severity

Stuttering severity was determined using the third edition of the Stuttering Severity Index (SSI) (Riley, 1994). AWS were video recorded, within one day of the acoustic and physiological data recording (described below), in a sound studio as they produced approximately five minutes of a monologue and read a passage aloud. Two certified speech-language pathologists, highly experienced with rating disfluent speech, transcribed each participant’s monologue and oral reading sample, and identified and rated stuttered events. If disagreement between the raters occurred they viewed the specific discrepancy as many times as necessary and conferred until an agreement was reached. For measures of reliability, the videotapes were re-evaluated six months later by two other experienced judges. Discrepancies between the second set of judges were resolved in a similar fashion. Video data were lost for one participant which made it impossible to estimate stuttering severity. Overall SSI ratings for the subjects used in the present study ranged from 16-35, with a mean rating of 25.5.

Acoustic/Physiological Data Recording

Participants were seated in a sound treated room and synchronous recordings of orofacial movement, chest wall circumference, and speech acoustics were made as they engaged in a variety of speaking tasks. Data collection sweeps were 30 seconds long and each sweep corresponded to a particular task. Typically, a single experimental session contained 16-20 sweeps. Motions of the upper lip, lower lip, mandible, and tongue blade were captured using a Carstens AG100 Electromagnetic Articulograph. Rib cage and abdominal circumference was recorded using an Ambulatory Monitoring Respitrace system. As this study is focused on
acoustic characteristics of speech, further details regarding the collection of articulatory and respiratory kinematic data are not provided. A Shure M93 miniature condenser microphone was positioned on the AG100 headpiece 7.5 cm from the mouth. The audio signal was amplified using a high quality preamplifier (Marenius MM-3300) and then digitized at a sample rate of 16 KHz using a custom sound board associated with the Carstens AG100.

Data Analyses

All analyses were performed on the nonsense phrase “a bad daba” (/ə bædæbə/) produced under two different speech rate conditions; casual and fast (self-selected to be twice the conversational speech rate). Subjects were instructed to place equal stress on the word “bad” and the first syllable of the word “daba.” For each of the two rate conditions, only perceptually fluent, well-articulated productions were included. The goal was to generate data for approximately 10 replicates from each subject within each test condition.

Decisions about the fluency and articulation of the samples were made initially by the student researcher. A second listener who is a licensed Speech Language Pathologist with substantial experience evaluating stuttered speech then verified the fluency of the tokens used in the study.

Utterance Duration/Speech Rate

The duration of each token was measured using a combined waveform-spectrogram display of TF32, an acoustic analysis software package (Milenkovic, 2000). These values were used to (1) verify that the different rate conditions actually resulted in measurable differences in
speech rate, and (2) to compare the degree to which the stuttering and non-stuttering groups altered speech rate.

First and Second Formant Transitions

Acoustic analysis focused on the word /bæd/. This word was selected because the transition from the bilabial plosive to the low front vowel typically results in marked formant transitions in both $F_1$ and $F_2$. These transitions can be observed in Figure 4, which shows a waveform and wide-band spectrogram of the test utterance drawn from the subject pool along with the key acoustic measures. The duration of the vowel /æ/ within the word “bad” will be determined using the waveform-spectrogram display of TF32 (see above). Vowel onset is defined as the first glottal pulse after the release burst of the preceding stop consonant (Robb & Blomgren, 1997). Vowel offset is defined as the final glottal pulse associated with the vowel just prior to the closure interval of the following stop consonant.

---

2 It is recognized that the speech stimuli used represents a narrow phonetic context and, under ideal circumstances, a greater range of consonant-vowel combinations would be included. However, the proposed study makes use of existing data that were collected for other purposes.
The first (F₁) and second (F₂) formant histories were estimated for the vowel using a multistep process involving SpeechTool, a locally developed software package (Gayvert & Hillenbrand, 2000). The LPEdit function was used to generate a linear predictive coding (LPC)–based spectrum at 10-msec intervals for the entire duration of the vowel. The spectral peaks were automatically identified by the software program at each interval. The time history of the spectral peaks was then overlaid on an LPC-based spectrogram. The peaks that corresponded to the F₁ and F₂ regions in the spectrogram were then extracted (see Figure 4).

The F₁ and F₂ transitions were determined separately. Transition onset was defined as the formant frequency at vowel onset. Transition offset was determined using the method described by Weismer, Kent, Hodge and Martin (1988), which defines a transition offset as a less than 20 Hz change over a 20 msec duration. This was an algorithmic procedure performed using a custom written, interactive Matlab program. The program identified all occurrences of the 20 Hz over 20 msec criterion and plotted their locations on the F₁ and F₂ time histories for visual inspection and selection. The earliest occurrence was selected as the transition offset in the vast
majority of tokens. However, in some cases, the first occurrence was not used because visual inspection indicated the formant transition continued after that time point. These were typically judged to be the consequence of an error in the formant tracking procedure. This was judged to have occurred on less than five percent of the tokens.

Once transition onset and offset were determined, $F_1$ and $F_2$ transition duration and transition extent were determined. Formant transition rate was calculated by taking the quotient of transition extent and duration for each formant.

**Statistical Analyses**

Five multi-level mixed-effects regression models were used to assess the effect of diagnostic category (i.e. AWS vs. NFS) and speech rate condition (i.e. casual vs. fast speech rate) on each of the dependent measures while controlling for repeated (correlated) observations obtained within subjects. Diagnostic category and speech rate conditions were treated as fixed factors. Observations were considered nested within conditions, and trials were nested within individual participants within this analytic model. Main and interaction effects were tested. To evaluate the role of measured speech rate on the results, each statistical model was generated with and without measured speech rate as a covariate.

Overall model significance was evaluated using the Wald Chi-Square statistic. Each of the five models, as well as the main effects and interaction terms were assessed using a $p$-value of 0.01 to control for an overall experiment-wide $p$-value of 0.05.
CHAPTER IV

RESULTS

A total of 1093 observations were made across the pool of 54 subjects. Although the original plan was to include at least 10 replicates per subject and condition, the total number of observations analyzed for each subject within each speech condition ranged from 5-12 replicates. This was because many of the replicates had to be excluded from analysis due to disfluency, misarticulation, diminished recording quality, or because it was difficult to discern a clear formant transition either by the formant extraction program or by the examiner. The last of these exclusions occurred most frequently for the $F_2$ transition.

The means and standard deviations for each of the dependent measures organized by diagnostic category and speech condition are summarized in Table 1. Generally, speech rates ranged from 4-5 syllables/sec for the casual speech condition and 6-7 syllables/sec for the fast speech condition. Typical $F_1$ and $F_2$ transition durations were 30-50 msec, while $F_1$ and $F_2$ formant transition rates ranged from 4-5 Hz/msec. These values are consistent with previous reports in the literature for similar consonant-vowel transitions (Zebrowksi, Conture, and Cudahy, 1985).
Table 1. Means and Standard Deviations for each measure based upon diagnostic category and speech condition.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Speech Rate</th>
<th>NFS</th>
<th>AWS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speech Rate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(syllables/sec)</td>
<td>Casual</td>
<td>4.72</td>
<td>0.53</td>
</tr>
<tr>
<td></td>
<td>Fast</td>
<td>6.44</td>
<td>0.87</td>
</tr>
<tr>
<td>F1 Duration (msec)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Casual</td>
<td>35.27</td>
<td>17.24</td>
<td>40.29</td>
</tr>
<tr>
<td>Fast</td>
<td>33.80</td>
<td>15.68</td>
<td>37.79</td>
</tr>
<tr>
<td>F1 Rate (Hz/msec)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Casual</td>
<td>3.99</td>
<td>2.09</td>
<td>4.14</td>
</tr>
<tr>
<td>Fast</td>
<td>4.49</td>
<td>2.16</td>
<td>4.24</td>
</tr>
<tr>
<td>F2 Duration (msec)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Casual</td>
<td>43.08</td>
<td>22.03</td>
<td>45.29</td>
</tr>
<tr>
<td>Fast</td>
<td>39.32</td>
<td>19.03</td>
<td>45.26</td>
</tr>
<tr>
<td>F2 Rate (Hz/msec)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Casual</td>
<td>4.45</td>
<td>1.98</td>
<td>4.63</td>
</tr>
<tr>
<td>Fast</td>
<td>4.83</td>
<td>1.91</td>
<td>5.13</td>
</tr>
</tbody>
</table>

Wald Chi-Square test statistics ($\chi^2$) and $p$-values for the overall mixed-effects regression models are shown in Table 2. The models were run twice; once with overall speech rate as a covariate, and again with overall speech rate excluded. It was determined from the results that including speech rate as a covariate did not markedly change the overall outcomes. Due to this observation, those results will not be reported. Details regarding each of the mixed-effects models can be found in Appendix B. It can be seen that the overall model fit was only significant for F1 transition rate and overall speech rate. Model fits for F1 and F2 duration and F2 rate approached, but did not reach significance. However, for many of the measures, trends in the data were observed. Therefore, results for all dependent measures will be reported below.
Table 2. Overall model values for Wald Chi-Square ($\chi^2$) statistic which were assessed based on a $p$-value of 0.01 to control for an overall experiment-wide $p$-value of 0.05.

<table>
<thead>
<tr>
<th>Model</th>
<th>$\chi^2$ (3)</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speech Rate</td>
<td>3310.47</td>
<td>0.00005</td>
</tr>
<tr>
<td>F1 Duration</td>
<td>7.18</td>
<td>0.067</td>
</tr>
<tr>
<td>F2 Duration</td>
<td>10.21</td>
<td>0.017</td>
</tr>
<tr>
<td>F1 Rate</td>
<td>11.56</td>
<td>0.009</td>
</tr>
<tr>
<td>F2 Rate</td>
<td>10.79</td>
<td>0.013</td>
</tr>
</tbody>
</table>

Speech Rate

Results for speech rate, organized by speaker group and speech condition, are shown in Figure 5. Both speaker groups markedly increased overall speech rate in the fast speaking condition. This observation was confirmed by a significant main effect for the speech condition ($z = 42.64, p < 0.00005$). There was no significant main effect for speaker group ($z = -1.08, p = 0.279$) nor was there an interaction between speaker group and speaking condition ($z = -1.16, p = 0.247$), suggesting that the two speaker groups performed the speech task at similar rates for both casual and fast rate conditions. In summary, AWS and NFS appeared to have produced both casual and fast utterances with statistically similar speech rates.
Figure 5. Mean and standard error of speech rate for the normally fluent speakers (NFS) and the adults who stutter (AWS) across the casual and fast (as labeled) rate conditions.

\[ F_1 \] Transition Duration

Results for \[ F_1 \] transition duration are shown in Figure 6. Recall that the overall model fit for this measure was not significant (\( \chi^2 = 7.18, p = 0.067 \)). However, there appears to be a trend for AWS to exhibit longer \[ F_1 \] transition durations overall compared to NFS. Further, fast speech tends to have a shorter duration relative to the casual speech condition for NFS. That being said, neither trend observed reached statistical significance (\( z = 1.80, p = 0.072; \) \( z = -1.16, p = 0.248 \) respectively). Furthermore, there was no statistically significant interaction between group and speech rate (\( z = -0.41, p = 0.684 \)).
Figure 6. Mean and standard error of first formant transition duration for the normally fluent speakers (NFS) and the adults who stutter (AWS) across the casual and fast (as labeled) rate conditions.

$F_2$ Transition Duration

See Figure 7 for results of $F_2$ transition duration analyses. Recall that when adjusted for multiple comparisons, the overall model did not reach significance ($\chi^2 = 10.21, p = 0.017$). However, it can be seen in Figure 7 that AWS tended to have longer $F_2$ durations compared to the NFS and that the two groups showed a different pattern of duration change across the two speech conditions. However, there was not a significant main effect for diagnostic group ($z = 0.91, p = 0.364$), nor was the interaction between group and speech condition significant ($z = 1.92, p = 0.055$). Like the overall model fit, the main effect for speech condition approached significance ($z = -2.6, p = 0.01$). Generally, it can be seen that although a trend exists for $F_2$ duration to be longer for AWS than NFS, these results were not statistically significant.
Figure 7. Mean and standard error of second formant transition duration for the normally fluent speakers (NFS) and the adults who stutter (AWS) across the casual and fast (as labeled) rate conditions.

\[ F_1 Transition Rate \]

Results for measures of \( F_1 \) transition rate can be found in Figure 8. For this measure, the overall model was significant when adjusted for multiple comparisons \( (\chi^2 = 11.56, p = 0.009) \). There was no significant main-effect for diagnostic category \( (z = 0.43, p = 0.67) \) indicating that \( F_1 \) rate was similar for the two groups. There was a significant main effect of speech condition indicating that \( F_1 \) rate was different for casual and fast rate conditions \( (z = 3.37, p = 0.001) \).

Specifically, it appears that transition rates are larger for the fast rate condition. However, in Figure 8, it appears that this pattern was only present for the NFS group. AWS did not make a commensurate increase adjustment of \( F_1 \) transition rate with an increase in overall speech rate. This observation was only a trend as the interaction between the group and speech rate was not
significant ($z = -2.0, p = 0.045$). In summary, AWS have a tendency to produce slower $F_1$ transition rates for fast speech compared to NFS, although not statistically significant.

![Figure 8](image_url)

**Figure 8.** Mean and standard error of first formant transition rate for the normally fluent speakers (NFS) and the adults who stutter (AWS) across the casual and fast (as labeled) rate conditions.

### $F_2$ Transition Rate

Results for $F_2$ transition rate can be found in Figure 9. As stated previously, the overall model did not quite reach statistical significance when making adjustments for multiple comparisons ($X^2 = 10.79, p = 0.013$). The main effect for diagnostic group was not significant ($z = 0.50, p = 0.617$) nor was the interaction between group and speech condition ($z = 0.52, p = 0.602$). The main effect for speech condition approached significance ($z = 2.08, p = 0.037$) indicating a trend for the $F_2$ transition rate to be greater for the fast speech rate condition.
Figure 9. Mean and standard error of second formant transition rates for the normally fluent speakers (NFS) and the adults who stutter (AWS) across the casual and fast (as labeled) rate conditions.
CHAPTER V

DISCUSSION

Civier et al. (2010) described a model of stuttering based on the DIVA computational model of speech production. Specifically, the model hypothesizes that, due to a problem with feedforward motor control, persons who stutter rely more heavily on a control strategy based on auditory feedback. A limitation of this type of control strategy is that there are inherent time lags needed for auditory information to reach the central nervous system and then for any corrective movements to be made. A feedback based strategy would predict subtle acoustic anomalies which would be observable in the fluent speech of AWS. Specifically, speech events characterized by fast acoustic transitions such as those observed in consonant-vowel transitions would be slower than those of NFS speakers who employ an open loop feedforward control strategy. The primary purpose of this study was to test the over-reliance on feedback hypothesis by examining $F_1$ and $F_2$ variations associated with a stop consonant-vowel transition in a test utterance produced at casual and fast speech rates by a group of AWS and NFS. The hypothesis would predict a number of differences between the two speaker groups. First, it would be a likely possibility that overall measured speech rates for AWS should be lower than NFS for both casual and fast speaking conditions as a compensation for the feedback based strategy. Second, it would be expected that AWS would exhibit longer $F_1$ and $F_2$ transition durations relative to NFS. Third, rate of $F_1$ and $F_2$ transitions rates should be lower for AWS. Finally, differences between the two speaker groups should be greater for the fast speaking condition.

In general, findings from the current study do not provide strong support for the overreliance on feedback hypothesis. Many of the observed differences between the two groups
were small and did not reach statistical significance. However, there were a number of non-significant trends in the data that were suggestive that the feedback hypothesis may be a possible cause of stuttering. These will be described below.

Unlike other studies of formant transitions in the stuttering population, the current study specifically manipulated speech rate, requiring the speaker groups to produce the test utterance at self-selected casual and fast speech conditions. The AWS produced both casual and fast speech rates which were comparable to NFS. This is in contrast to some previous studies which found that AWS exhibit slightly lower speech rates in comparison with NFS (Robb & Blomgren, 1997). This finding is also in disagreement with the hypothesis being tested in the present study. It would be expected that, if an over-reliance on feedback control is present, AWS would be limited in the rate with which they produce speech, at a casual rate but especially at the fast speech condition. Results from the current study show that AWS speech rate modifications were not statistically different from NFS. This result could be due to the self-selected rate modification not being enough of a challenge to the motor system of the AWS. It is plausible that by utilizing a more consistent control of speech rate to amply tax the system of AWS, a difference may be seen between the populations.

There was a non-significant trend for AWS to exhibit longer $F_1$ and $F_2$ transition durations than the NFS group. Previous research has often neglected to examine $F_1$, which is known to exhibit marked transitions during consonant-vowel boundaries. Blomgren, Robb, and Chen (1998) did evaluate $F_1$ values in AWS and NFS. Their measures, however, focused on vowel steady states rather than formant transition duration. As a result, comparable results for $F_1$ transitions are not available in the literature. $F_2$ transition durations have been measured previously, and these studies reported that $F_2$ transitions for people who stutter differ from those
of NFS, but that the difference is not significant, a result which is supported by the present study (Yaruss & Conture, 1993; Zebrowski, Conture, and Cudahy, 1985). AWS demonstrated a tendency for longer formant transition durations, which is as would be expected if they are relying more heavily on a feedback control system. Additionally, AWS showed a tendency to produce longer F₂ transition durations for fast speech compared to casual, a finding which is the opposite of that found for NFS. These trends indicate that as AWS increased overall rate of speech, their F₂ transition durations actually got longer.

Analyses from the current study indicated a non-significant trend for AWS to have lower F₁ and F₂ rates for the casual speech condition, while having higher F₁ and F₂ rates for the fast speech condition compared to NFS. As mentioned previously, there is literature examining F₁ transitions is lacking; however F₂ transition data is much more prevalent. The work of Chang, Ohde, and Conture (2002) showed no evidence for a difference in F₂ transition rates between children who do and do not stutter. This finding is supported by the present study, as no significant differences were found between AWS and NFS. Results from the present study indicated that there is a lack of increase in F₁ rate at the fast speech condition for the AWS. On the other hand, F₂ transition rate differences between speech conditions for AWS were more typical of those seen by the NFS group. The hypothesis predicts that formant transition rates should be slower for AWS as they are thought to rely more heavily on feedback information, and that this would be more evident for the fast speech condition. Based on the trends observed in the present study, this hypothesis is somewhat supported by the F₁ transition rate findings; however it is not supported by the F₂ rate data.

Although some trends that support the hypothesis were observed in the present study, only F₁ transition rate and overall speech rate were found to be statistically significant. The
other differences noted were not statistically significant when $p$-values were adjusted for multiple comparisons.

Study Limitations and Future Directions

There are a number of limitations associated with the current study. It is somewhat difficult to make direct comparisons of these findings with previous studies of formant transitions in persons who stutter due to methodological differences. The approach used for identifying formant transitions in this study, while common in the motor speech disorders literature (Weismer et al., 1988) has not been used previously for this type of research.

Additionally, the narrow phonetic context utilized by this study makes generalizations to other speech elements that contain formant transitions difficult. This could not be avoided in the current study given that it exploited data which were collected for other purposes. Although the formant transition between the initial bilabial plosive /b/ into the subsequent vowel /æ/ is rapid and large in size, it would be useful to examine other phonetic contexts where significant formant transitions occur (i.e. diphthongs, liquids, glides).

Another limitation to the present study is that speech rate was a self-selected modification. This variable was controlled for and not found to be statistically significant, however the trend for AWS to speak more slowly in both speech rate conditions could indicate that more consistent control could impact the results. Future research should consider additional control of speech rate to eliminate any chance of a self-selected slowed overall speech-rate.
Appendix A

Human Subjects and Institutional Review Board
Date: February 16, 2015

To: Stephen Tasko, Principal Investigator
   Kaitlin Arnold, Student Investigator

From: Amy Naugle, Ph.D., Chair

Re: HSIRB Project Number 11-02-35

This letter will serve as confirmation that the change to your research project titled “Neuromotor Basis of Stuttering: Data Analysis” requested in your memo received February 13, 2015 (to remove student investigators Kaitlin Abbs, Danielle Gage, Hannah Hodges, Megan Pachesny, and Rachel Whitney) has been approved by the Human Subjects Institutional Review Board.

The conditions and the duration of this approval are specified in the Policies of Western Michigan University.

Please note that you may only conduct this research exactly in the form it was approved. You must seek specific board approval for any changes in this project. You must also seek reapproval if the project extends beyond the termination date noted below. In addition if there are any unanticipated adverse reactions or unanticipated events associated with the conduct of this research, you should immediately suspend the project and contact the Chair of the HSIRB for consultation.

The Board wishes you success in the pursuit of your research goals.

Approval Termination: February 15, 2016
Appendix B

Mixed-Effects Model Details for all Measures
### Mixed-Effects Model for Overall Speech Rate

<table>
<thead>
<tr>
<th></th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>z-score</th>
<th>p-value</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main-Effect: diagnostic category</td>
<td>-0.19</td>
<td>0.17</td>
<td>-1.08</td>
<td>0.28</td>
<td>-0.53 - 0.15</td>
</tr>
<tr>
<td>Main-Effect: speech condition</td>
<td>1.75</td>
<td>0.04</td>
<td>42.64</td>
<td>0.000</td>
<td>1.67 - 1.83</td>
</tr>
<tr>
<td>Interaction: diagnostic category/speech condition</td>
<td>-0.07</td>
<td>0.06</td>
<td>-1.16</td>
<td>0.25</td>
<td>-0.19 - 0.05</td>
</tr>
</tbody>
</table>

### Mixed-Effects Model for F1 Transition Duration

<table>
<thead>
<tr>
<th></th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>z-score</th>
<th>p-value</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main-Effect: diagnostic category</td>
<td>5.02</td>
<td>2.79</td>
<td>1.80</td>
<td>0.07</td>
<td>-0.45 - 10.49</td>
</tr>
<tr>
<td>Main-Effect: speech condition</td>
<td>-1.53</td>
<td>1.33</td>
<td>-1.16</td>
<td>0.25</td>
<td>-4.13 - 1.07</td>
</tr>
<tr>
<td>Interaction: diagnostic category/speech condition</td>
<td>-0.79</td>
<td>1.93</td>
<td>-0.41</td>
<td>0.68</td>
<td>-4.58 - 3.00</td>
</tr>
</tbody>
</table>

### Mixed-Effects Model for F2 Transition Duration

<table>
<thead>
<tr>
<th></th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>z-score</th>
<th>p-value</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main-Effect: diagnostic category</td>
<td>-02.60</td>
<td>2.87</td>
<td>0.91</td>
<td>0.36</td>
<td>-3.01 - 8.22</td>
</tr>
<tr>
<td>Main-Effect: speech condition</td>
<td>-4.14</td>
<td>1.59</td>
<td>-2.60</td>
<td>0.01</td>
<td>-7.25 - 1.02</td>
</tr>
<tr>
<td>Interaction: diagnostic category/speech condition</td>
<td>4.72</td>
<td>2.46</td>
<td>1.92</td>
<td>0.06</td>
<td>-0.1 - 9.54</td>
</tr>
</tbody>
</table>
### Mixed-Effects Model for F1 Transition Rate

<table>
<thead>
<tr>
<th></th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>z-score</th>
<th>p-value</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main-Effect: diagnostic category</td>
<td>0.15</td>
<td>0.35</td>
<td>0.43</td>
<td>0.67</td>
<td>-0.54 to 0.84</td>
</tr>
<tr>
<td>Main-Effect: speech condition</td>
<td>0.48</td>
<td>0.14</td>
<td>3.37</td>
<td>0.001</td>
<td>0.201 to 0.76</td>
</tr>
<tr>
<td>Interaction: diagnostic category/speech condition</td>
<td>-0.42</td>
<td>-0.21</td>
<td>-2.00</td>
<td>0.05</td>
<td>-0.82 to -0.01</td>
</tr>
</tbody>
</table>

### Mixed-Effects Model for F2 Transition Rate

<table>
<thead>
<tr>
<th></th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>z-score</th>
<th>p-value</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main-Effect: diagnostic category</td>
<td>0.15</td>
<td>0.3</td>
<td>0.50</td>
<td>0.62</td>
<td>-0.44 to 0.74</td>
</tr>
<tr>
<td>Main-Effect: speech condition</td>
<td>0.32</td>
<td>0.15</td>
<td>2.08</td>
<td>0.04</td>
<td>0.02 to 0.62</td>
</tr>
<tr>
<td>Interaction: diagnostic category/speech condition</td>
<td>0.12</td>
<td>0.24</td>
<td>0.52</td>
<td>0.60</td>
<td>-0.34 to 0.59</td>
</tr>
</tbody>
</table>
REFERENCES


