2016

Phonological Encoding of Medial Vowels in Adults Who Stutter

Allison Elizabeth Jacobs
Louisiana State University and Agricultural and Mechanical College, ajaco19@lsu.edu

PHONOLOGICAL ENCODING OF MEDIAL VOWELS IN ADULTS WHO STUTTER

A Thesis

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Master of Arts

in

The Department of Communication Sciences and Disorders

by
Allison E. Jacobs
B.A., University of Colorado, 2013
May 2016

ACKNOWLEDGEMENTS

I would like to thank my major professor, Dr. Geoffrey A. Coalson, for guiding me throughout this process and sharing his wealth of research experience and academic knowledge.

Special thanks also go to Dr. Yunjung Kim and Dr. Paul Hoffman for serving as members on my committee for this thesis.

Finally, I would like to express my sincere gratitude to Lauren Davis Paxton, Hailey Guillot, Amanda Kuylen, Laura Fuller, Joseph Moeglin, and Francesca Tusa for the extensive time and effort they devoted to helping me carry out my research.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Previous data suggest the metrical properties of a word may influence the time course of phonological encoding, particularly in adults who stutter. The purpose of the present study is to examine phonological and metrical encoding skills in fluent and non-fluent adults, in particular the medial stress-bearing vowel. Investigators used a silent phoneme monitoring paradigm to assess reaction times for all phonemes within nonword CVCCVC stimuli. This paradigm required participants to manually identify target phonemes within a nonword to further isolate the level of phonological encoding from other processes. Eight participants were exposed to stimuli with initial-stress, and eight participants were exposed to stimuli with non-initial stress. Both groups demonstrated increased monitoring latencies for the initial vowel, regardless of initial or non-initial stress. However, results did not yield any significant between-group latencies or false negative errors. Participants from both groups demonstrated increased post-trial error rates compared with those from previous studies with similar methodology, suggesting the task of auditory identification of consonants and vowels in isolation may have been more challenging for both talker groups than identification of consonants alone. Together, these preliminary data suggest that AWS and AWNS demonstrate similar efficiency when encoding the medial, stress-bearing vowel.

CHAPTER I: INTRODUCTION

Stuttering is a disorder of speech in which an intended message is interrupted by involuntary stoppages such as single-sound repetitions (e.g., m-m-m-magnet), prolongations (e.g., mmmmagnet), or blocks (e.g., ---magnet) (Bloodstein & Bernstein Ratner, 2008). Developmental stuttering has a prevalence of approximately 1% and a lifetime incidence of more than 5% within the population (Yairi & Ambrose, 2013). Alongside experiencing communicative difficulties, people who stutter are often ascribed negative stereotypes, and frequently suffer from diminished social-emotional functioning and overall quality of life (Craig, Blumgart, & Tran, 2009). In addition, many speech-language pathologists report low levels of confidence and competence regarding their effectiveness in providing stuttering therapy (e.g., Coalson, Byrd, & Rives, in press; Cooper & Cooper, 1985, 1996; Gabel, 2014; St. Louis & Durrenberger, 1993). To provide effective therapy for this clinical population, it is critical to first identify the underlying factors which may provoke moments of disfluent speech.

Research suggests stuttering may occur due to difficulties planning speech prior to its execution (see Postma & Kolk, 1993; Sasisekaran, De Nil, Smyth, & Johnson, 2006; Wolk, Edwards, & Conture, 1993; cf. Nippold, 2002). Several theoretical models implicate deficits or differences in phonological encoding as the critical factor that precedes stuttered speech (e.g., Wingate, 1988; Postma & Kolk, 1993). Previous studies have demonstrated that in typical speech, sounds within a word are assigned to syllabic frames prior to production, and certain sounds within words are not fully prepared prior to production. One experimental paradigm, the silent phoneme monitoring task, has yielded data which support that consonants within target words are assembled from onset to coda at incrementally slower rates in typically fluent adults (e.g., Wheeldon & Levelt, 1995; Wheeldon & Morgan, 2002). Silent phoneme monitoring tasks have also been applied to individuals who stutter to examine potential differences in speech planning (e.g., Coalson & Byrd, 2015; Sasisekaran et al., 2006; Sasisekaran, Brady, & Stein, 2013) and, in support of theoretical models of stuttering, delayed phonological encoding was observed in participants who stutter.

Although these studies provide preliminary evidence that individuals who stutter may experience phonological encoding differences, all are characterized by one notable limitation – a lack of vowels as silent phoneme monitoring targets. Wingate's (1988) model of stuttering suggests that the intervening vowel is the critical point of breakdown in speech planning for those who stutter. Without knowledge regarding phonological encoding of vowels, validation of theoretical models of stuttering remains incomplete. Therefore, the present study includes vowels as targets to examine speech planning abilities in adults who stutter (AWS) and those who do not (AWNS). Findings from this study may allow for reinforcement or modification of existing theoretical models of stuttering, and may ultimately contribute to more effective methods of clinical intervention.

Review of the literature

Descriptive and psycholinguistic research implicates subtle weaknesses in phonological encoding in AWS relative to typically fluent adults. Ample data indicate the frequent co-occurrence of stuttering with phonological disorders in children (Arndt & Healey, 2001) and at disproportionately higher levels than non-stuttering children (Blood, Ridenour, Qualls, & Hammer, 2003). More recent longitudinal studies have implicated phonological encoding difficulties as one potential predictive factor for the persistence of stuttering beyond childhood (e.g., Spencer & Weber-Fox, 2014; Ambrose, Yairi, Loucks, Seery, & Throneburg, 2015). Closer examination of the nature of the breakdown during speech planning in individuals who stutter may provide specification to these data, how phonological encoding processes may deviate in individuals who stutter relative to typically fluent speakers, and how these deviations may be linked to actual moments of stuttered speech.

One model of phonological encoding proposed by Levelt and colleagues (Levelt, 1989; Levelt et al., 1999) defines phonological encoding as "the processes involved in retrieving or building a phonetic or articulatory plan for each word and the utterance as a whole" (p. 12). According to this model, phonological encoding is compartmentalized into three separate stages: (a) the generation of sound segments which comprise a word (i.e., segmental encoding), (b) syllabic stress and syllable boundary assignment (i.e., metrical encoding), and (c) integration of these sound segments to their appropriate locations within the framework of the word (i.e., syllabification). To examine the potential deficits involved in stuttered speech, it is crucial to first describe the processes underlying typical speech. Therefore, the present study examines how efficiently phonological encoding occurs in the absence of speech production in AWNS, and to what degree the specific phonological properties of a word – such as medial vowels – may facilitate or impede phonological encoding in AWS.

Phonological encoding in typically fluent speakers

Levelt and colleagues (Levelt, Roelofs, & Meyers, 1999) proposed a theory of speech production which accounts for all stages of speech, from the mental preparation of a concept to overt initiation of the articulatory movements required to verbally express it. According to this model, a variety of processes occur prior to the emergence of speech itself. First, a particular lemma, or mental concept, is activated according to the speaker's intended meaning (e.g., "four-legged canine house pets"). The speaker then selects the intended word form, (e.g., "dog"), which is followed by morphological encoding (e.g. + "-s"), phonological encoding (e.g. /d-ɑ-g-z/), phonetic encoding (e.g., [dɑgz]), and, finally, the initiation and execution of the motor speech plan. The process which occurs just prior to phonetic encoding - phonological encoding - must occur within a relatively brief time frame with minimal errors in order to facilitate fluent speech production (e.g., Hartsuiker, 2007; Indefrey & Levelt, 2004).

According to Levelt and colleagues (1999), as phonological encoding takes place, a speaker uses self-monitoring via auditory comprehension mechanisms to ensure the integrity of the intended message as he or she is producing it. In addition to the external monitoring system, speakers have the ability to internally monitor a representation of the speech plan before initiating the physical act of speech. This pre-articulatory monitoring during the encoding process allows the speaker to detect output errors in the speech plan before they are produced. Measurement of a speaker's ability to silently monitor phonemes prior to production provides a general estimation of the speed and accuracy of phonological encoding in typically fluent speakers.

Silent phoneme monitoring tasks in typically fluent speakers

A number of experimental paradigms have been developed with the intent of examining errors made during speech production. Many of these have involved either (a) analysis of spontaneous or evoked errors, or (b) latencies and/or errors in picture naming tasks (for review, see Levelt, 1999). While both of these methods offer valuable information about a number of aspects of speech production, the level of breakdown in those who stutter remains unclear, as semantic and syntactic selection, phonological encoding, rapid syllabification, and articulatory preparation may have all served as influential factors in these studies. That is, because overt speech is the end-product of a rapid cascade of events, any or all of these factors may have contributed to differences observed in previous experiments seeking to examine the role of a single, isolated level of processing.

In contrast to previous experimental methods, the *silent phoneme monitoring task* isolates the level of phonological encoding from other processes of speech preparation and overt speech. During a silent phoneme monitoring task, a participant provides a non-verbal response via a button press to the presence or absence of a pre-specified phoneme within a target word. For example, the participant is prompted to identify a particular sound, such as /g/, in the upcoming word. The participant is then presented with a picture of a magnet, and required to press the 'Yes' button as quickly and accurately as possible. Phonemes which have been used as targets in original silent phoneme monitoring tasks conducted by Wheeldon and colleagues (Wheeldon & Levelt, 1995; Wheeldon & Morgan, 2002) included all consonants within a word. That is, for the word "magnet," the participant was required to respond to the initial consonant (C_1, e.g., /m/), second consonant (C_2, e.g., /g/), third consonant (C_3, e.g., /n/), and fourth consonant (C_4, e.g., /t/).

Wheeldon and Levelt (1995) were the first to use this experimental methodology to quantify the time frame during which phonological encoding occurs. In their study, bilingual Dutch participants silently translated aurally presented words from English to Dutch, and internally monitored their own translations for a pre-determined consonant target. Investigators measured monitoring latencies with bisyllabic stimuli for first and second syllable-initial segments, and in non-initial stress stimuli. Investigators found that subjects took less time to respond to word onset targets as compared to second syllable targets, regardless of syllabic stress. In addition, monitoring latencies suggested the first syllable was completely encoded before encoding of the second syllable began. These findings, based upon phonological encoding of consonants only, suggest that phonemes are assigned to their appropriate frames in a left-to-right manner, with encoding of segmental and metrical information within the first syllable occurring prior to encoding of the second syllable.

Wheeldon and Morgan (2002) conducted multiple experiments to replicate and extend the findings of Wheeldon and Levelt (1995). Participants (*n*=40) were British, monolingual, English-speaking adults. The aim of the study was to lend further clarity to the phonological representation a speaker uses during the prearticulatory stage, as well as the time frame in which it is generated. Similar to Wheeldon and Levelt, stimuli were comprised of bisyllabic words with CVCCVC patterns (i.e., consonant-vowel-consonant prior to the syllable boundary, consonant-vowel-consonant after the syllable boundary, e.g. /m-æ-g . n-ɛ-t/). In Experiment 1, investigators measured latencies for participants' silent monitoring of the four consonants in the CVCCVC words – C_1, C_2, C_3, and C_4. Prior to experimental testing, participants demonstrated their abilities to rapidly produce the stimuli without articulatory errors. Participants were then primed with the upcoming target consonant in isolation ("React now to the sound /tə/") and within words which represented the consonant in multiple target positions. Participants were required to monitor a series of words via response key for whether or not they contained the target consonant. Results from this study indicated that phonological assignment occurs in a serial manner from left-to-right, and differences are apparent at the syllable boundary between the consonants at the word medial position. In addition, significant latency differences were observed between C_1 (first consonant) and C_2 (second consonant), secondarily between C_2 and C_3 (the consonants on either side of the syllable boundary), and, finally, with the least difference between C_3 and C_4 (the final two consonants). These results were interpreted to suggest that not only does segmental encoding occur in a serial manner, but metrical properties may further delay the phonological encoding process. Specifically, increased latencies between C_1 and C_2 represent assignment of metrical stress on the intervening vowel for words with initial stress, and in the absence of initial stress, assignment of the metrical syllable boundary is encountered first and therefore requires additional processing time (see Jansma & Schiller, 2004; Schiller, 2005; Wheeldon & Levelt, 1995, p. 322).

Unlike Wheeldon and Levelt (1995), Wheeldon and Morgan (2002) directly tested the interpretation that delayed C_1 - C_2 latencies were due to assignment of the intervening, stress-bearing vowel. Experiment 3 of their study included participants' internal monitoring of the initial CVC syllable of bisyllabic stimuli (with initial stress) which were similar to those used by Wheeldon and Levelt, and measured the speed of retrieval and encoding at vowel and consonant levels within the initial CVC syllable of bisyllabic stimuli. Results indicated significant latency differences between the syllable onset (C_1) and intervening vowel (V_1) as well as V_1 and C_2, with C_1-V_1 differences greater than those between the intervening vowel and the coda. By directly examining the medial vowel, these data provide support for the interpretations of Wheeldon and Levelt regarding metrical encoding in typically fluent adults, and empirical foundation to examine theories of stuttering that suggest the intervening vowel requires longer to plan during syllabification for AWS due to difficulties with both segmental and metrical encoding.

Phonological encoding in individuals who stutter

Previous studies have examined self-monitoring as a function of prearticulatory speech error correction (Levelt et al., 1999). Postma and Kolk (1993) proposed a compatible theoretical model of stuttering to describe how this system may function in the presence of stuttered speech. In their model, known as the Covert Repair Hypothesis, repairs in the internal speech plan may disrupt overt, ongoing speech. A typical adult speaker produces 2-3 words per second, and a speech error about every 1,000 words uttered (Levelt, 1999). These errors themselves do not typically result in the disruption of speech; however, the subsequent repairs might (Postma & Kolk, 1993). A disruption of this nature may occur during overt speech if a speaker attempts to initiate a correction for an error anticipated on a subsequent segment. This conflict may result in a brief interruption in speech, which is the overt equivalent of the speaker's internal detection of a speech plan error and postponement of speech production until an accurate plan has been formulated. For a typical individual, these interruptions are brief, as the appropriate speech planning systems (e.g., morphological or phonological encoding) rapidly initiate a corrected speech plan, at which time the interruption ceases and the correct message is produced with only a brief disfluency of typical nature (e.g., "mac---um---magnet"). However, if the system is burdened with a disproportionately high amount of phonological processing deficits or phonemic errors, as may be the case in individuals who stutter, repairs may result in disfluencies characteristic of stuttered speech, such as single-sound repetitions, blocks, or sound prolongations (e.g., "m-m-m-magnet") as the speaker experiences delays while selecting the correct sequence of sounds. When a stuttered disfluency results from an attempted correction, the extent of its duration and severity depends largely upon the speed at which the error is detected, and the time taken thereafter to initiate and complete an appropriate correction. If the segmental and metrical properties of a word influence the speed and accuracy of its production, covert repairs may be subject to similar influences.

A theory posited by Wingate (1988) – the Fault-Line Hypothesis – claims that stuttered speech occurs due to delays when encoding the rhyme, or stressed syllable, prior to speech production. That is, individuals who stutter demonstrate greater difficulty with the blending of the initial consonant with the stress-bearing vowel prior to speech initiation. Wingate posits this may lead to a delay in the retrieval and encoding of the syllable rhyme during speech production, resulting in a divide, or "Fault-Line," at the point of syllable onset and rhyme integration. Wingate supported this notion with descriptive observations that stuttered speech in adults often occurs on the initial, stressed syllable of a word. These observations have been replicated in descriptive studies in children who stutter (e.g., Natke, Sandrieser, van Ark, Pietrowsky, & Kalveram, 2004) and adults who stutter (e.g., Natke, Grosser, Sandrieser, & Kalveram, 2002), but only two experimental investigations have used experimental measures to test the Fault-Line Hypothesis.

Wijnen and colleagues directly tested Wingate's (1988) claims using implicit naming tasks. In Wijnen and Boers (1994), nine AWS and nine AWNS received priming prior to an experimental task in order to generate pre-specified words which were semantically or phonetically-related to bisyllabic stimuli with initial stress. Results of the study indicated that for AWS, priming was only beneficial when a cue word shared both consonant and subsequent vowel in common with the target stimulus. These preliminary results from Wijnen and Boers' (1994) study were interpreted to support the hypothesis that the syllable rhyme presents the most salient problems for those who stutter. Burger and Wijnen (1999) attempted to replicate and extend the findings of Wijnen and Boers, with the addition of non-initial stress stimuli. Fifteen AWS and 15 AWNS participated in the experiment. Results indicated that inclusion of the medial vowel within the cue did not benefit AWS more so than AWNS, and performance did not differ between groups regardless of the presence or absence of initial stress. Thus, the authors found no substantial evidence to implicate phonological encoding of the stress-bearing portion of a syllable in individuals who stutter. However, the authors acknowledge that almost any perceived difference between AWS and AWNS could be reflective of baseline reaction time differences often observed in AWS (for review, see Smits-Banstra, 2010), and not necessarily the underlying process which an investigator might be attempting to measure (e.g., phonological encoding).

At a minimum, it seems plausible that the segmental and metrical properties of a word may influence latencies associated with successful phonological encoding, particularly in individuals who stutter. Byrd, Conture, and Ohde (2007) conducted an experiment which measured speech reaction times in children who stutter (CWS, n=26) and age-matched children who do not stutter (CWNS, n=26) to investigate the relationship between phonological maturity and stuttering. Tasks involved picture naming in three separate priming conditions – neutral (e.g., no priming), holistic (e.g., priming at the syllable or whole-word level), and incremental (e.g., priming a word via its sequential sounds). These conditions were appropriate constructs of measurement for the maturity of the children's phonological plans, as age-appropriate CWNS typically experience a transition from the use of holistic processing to incremental processing during childhood. Results for CWNS were consistent with investigators' predictions, in that typical children shifted from being faster in the holistic priming condition to being faster in the incremental priming condition. CWS, however, only demonstrated evidence of this re-organization at age 5 and significantly later than their CWNS age-matched peers. Researchers concluded that this key difference indicates delayed phonological skills in children who stutter. These differences, though they may be subtle, can create difficulties in the planning and maintenance of fluent speech. However, as in previously noted paradigms, this experiment measured reaction times using subjects' overt speech, which by definition fails to isolate phonological encoding from the motor processes involved in speech production. Therefore, in order to probe for results similar to those reported in this study, the use of another experimental paradigm, silent phoneme monitoring, may be beneficial.

Silent phoneme monitoring in individuals who stutter

Sasisekaran et al. (2006) conducted a study to measure phonological encoding in the silent speech plans of 10 AWS and 11 AWNS. Following a familiarization task, participants completed an overt naming task, a monitoring task for the detection of target phonemes while silently naming pictures, and a simple motor task in response to pure tone aural stimuli. Results indicated that AWS exhibited significantly longer latencies in phoneme monitoring compared to AWNS. However, there were no significant between-groups differences for response speed, picture naming accuracy, or simple motor tasks. Further, these reaction time data were dissimilar and longer than speakers' monitoring of auditory pure-tone sequences. These results suggest deficits at the level of phonological monitoring in AWS, and not deficits in baseline reaction time, or auditory or perceptual monitoring abilities.

A similar study conducted by Sasisekaran et al. (2013) investigated phonological encoding skills in CWS and CWNS. Participants included CWS (n=9) and age and sex-matched CWNS peers (n=9), all between 10 and 14 years of age. Participants completed familiarization tasks, then monitoring tasks in response to internally named pictures. Stimuli were comprised of 12 black and white line drawings depicting bisyllabic nouns (CVC(C)CVC, CVCCCV). All target stimuli were characterized by initial stress and target consonants at the onset and offset of each syllable. As demonstrated in previous experiments, CWS exhibited significantly longer times for identifying phonemes than their CWNS peers. Therefore, it is plausible that the same phonological encoding deficits present in CWS may persist into adulthood, continuing to cause disfluencies in AWS.

A recent study by Coalson and Byrd (2015) used nonword CVCCVC stimuli to examine the influence of the metrical properties of a word on phonological encoding in AWS and AWNS. Participants included 44 total adults, or 22 per experiment (11 AWNS and 11 AWS in each experiment). In Experiment 1, participants monitored nonwords with initial stress patterns for pre-specified phonemes. Stimuli and methodology in Experiment 2 were identical to Experiment 1, but included stimuli with non-initial stress. Coalson and Byrd found that latency values were not significantly different between talker groups while monitoring phonemes in initial-stress stimuli. However, for stimuli with non-initial stress, AWS exhibited significantly longer times when monitoring the consonant immediately following the syllable boundary (C_3). In addition, during silent phoneme monitoring tasks for non-initial stress stimuli, AWS demonstrated compromised accuracy when silently identifying sounds immediately prior to the syllable boundary (C_2), as well as errors in phoneme production and stress assignment during post-trial verbal response. These results further support that AWS may demonstrate significant differences with phonological encoding at the syllable boundary of words with non-initial stress compared to their AWNS peers, or, at minimum, both segmental and metrical information may influence the time course of phonological encoding for AWS. Results may also be interpreted to indicate that the presence or absence of a stress-bearing vowel in the initial syllable may not necessarily be prerequisite for delayed encoding between C_1 and V_1 to occur, as C_1 - V_1delays were observed in stimuli with initial stress (Experiment 1) and non-initial stress (Experiment 2). However, as in previous studies (Wheeldon & Levelt, 1995; Wheeldon & Morgan, 2002; Sasisekaran et al., 2006, 2013), these authors did not monitor the medial vowel. Therefore, further investigation is warranted to verify claims regarding the medial vowel in stressed and unstressed contexts, which serves as the foundation for the present study.

Purpose of the present study

Despite the aims of previous research to examine the relationship between speech planning and disfluencies, the information available is characterized by notable limitations. All studies except one by Coalson and Byrd (2015) were restricted in their use of homogenous word stress; for instance, there are multiple studies which describe moments of disfluency occurring on the first syllable of initial-stress stimuli. Perhaps more critically, while many studies have examined AWS' monitoring of consonants (Wheeldon & Morgan, 2002; Sasisekaran et al., 2013), none have measured latency times for the intervening vowels. This is significant because according to multiple theories (Postma & Kolk, 1993; Wingate, 1988), stuttered speech may be characterized by delayed encoding of the vowel and syllable stress. Without data which directly examine the speed and accuracy of vowel encoding in the presence and absence of syllabic stress, our understanding of phonological encoding in people who stutter is incomplete. This study, therefore, isolates the influence of segmental and metrical properties during the phonological encoding processes in those who stutter compared to their fluent peers, with specific focus on the initial, intervening vowel. This preliminary but novel data may lend further specificity to models which implicate phonological encoding deficits as a significant factor in stuttered speech. Thus, the present study examines two specific research questions:

1) Do AWS differ from AWNS during phonological encoding of initial medial vowels (V_1) in CVCCVC stimuli with initial stress?
2) Do AWS differ from AWNS during phonological encoding of initial medial vowels (V_1) in CVCCVC stimuli with non-initial stress?

CHAPTER II: METHODS

As proposed by Wingate (1988), AWS may be slower and less accurate than AWNS when assigning the initial, medial vowel prior to speech production, particularly if the medial vowel also carries syllabic stress. Thus, the purpose of the present study was to examine potential differences in speed and accuracy in assignment of the medial vowel to the initial syllable for AWS and AWNS during silent phoneme monitoring tasks. This study also served to replicate and extend data collected by Coalson and Byrd (2015) by using an identical paradigm and including vowels as well as consonants as target phonemes. Therefore, the methodology and analyses were identical to those of Coalson and Byrd (2015). The following will review these methods, and emphasize the key difference between the two studies – that is, inclusion of medial vowels in the initial stressed syllable (Experiment 1) and the initial unstressed syllable (Experiment 2).

Experiment 1: Initial Vowel in the Presence of Syllabic Stress

In accordance with Research Question 1, the purpose of Experiment 1 was to examine whether AWS differ from AWNS when encoding the medial vowel (V_1) with initial stress in CVCCVC stimuli. Participant inclusion criteria, talker group classification, intake tasks, procedures, and analyses for Experiment 1 were as follows:

Participants

Eight total participants (AWNS, *n* = 4, AWS, *n* = 4) participated in two 90-minute sessions. Subjects received compensation for their time and participation in the form of class credit or financial reimbursement. Following provision of informed consent, investigators administered a battery of pretests as described in Coalson and Byrd (2015) to determine whether or not a participant met the inclusion criteria to participate in the experimental portion of the study, which took place during the second session. Participants completed the following tasks: (a) self-report questionnaire, (b) hearing, vision, and handedness screening, (c) talker group classification tasks, (d) phonological processing subtests, and (e) identification of target phonemes in isolation.

Talker group classification

The two talker groups, AWS and AWNS, were matched for gender and age to the greatest extent possible. As detailed in Coalson & Byrd (2015), a 9-point stuttering severity scale (O'Brian, Packman, Onslow, & O'Brian, 2004) was used to rate participants' speech samples during conversational speech and a reading sample. Classification as AWS required: (a) self-identification as AWS with reported onset of stuttering prior to 7 years of age, and (b) a previous formal diagnosis of stuttering from a licensed speech-language pathologist. Classification as AWNS required: (a) self-identification as an individual who does not stutter, and (b) a rating of no higher than 1 (*no stuttering*) on the 9-point stuttering severity scale for both conversational and reading samples.

Participant exclusion criteria

Similar to Coalson & Byrd (2015), participants were excluded from the experimental portion of the study if they: (a) were unable to identify all of the target phonemes in isolation with 100% accuracy, (b) demonstrated mean baseline reaction times of two standard deviations below the mean of their talker group, (c) lacked native-like proficiency in Standard American English, or (d) failed to demonstrate adequate hearing or vision on either screening.

Identification of target phonemes in isolation

Prior to the experimental portion of the study, participants demonstrated 100% accuracy in auditory recognition of all target phonemes. First, participants took part in a brief familiarization task, during which each target phoneme was presented in isolation via headphones while the participant viewed the corresponding letter(s) on the computer screen. No response was required during the familiarization task. Immediately following the familiarization task, participants demonstrated auditory recognition of target consonants by facing away from the computer screen and verbally identifying the letter(s) which correspond with each phoneme or pointing to the correct orthographic representation of the vowel phoneme from a list of possibilities. To diminish ambiguity in participant responses to vowel targets, participants non-verbally indicated the correct target from a short list of vowel sounds extracted from "H(vowel)D" format (e.g., "hod" for /a/, "hed" for /ɛ/, etc.; see Hillenbrand, Getty, Clark, & Wheeler, 1995). All vowels were recorded in accordance with normative duration data (Hillenbrand et al., 1995) to ensure integrity of presentation adequate for accurate detection (see Table 1 below). In addition, all were presented aurally via headphones at 25 dB-SPL, which has been demonstrated to be adequate for accurate detection of potentially consonant targets (i.e., [f, v, z]; Pirello, Blumstein, & Kurowski, 1997). In the event that a participant produced the phoneme incorrectly after monitoring, the reaction time for the associated stimulus was excluded from analysis.

Table 1. Target vowel and range of duration.

Target Vowel	H-D Equivalent	Duration (mean)	Standard Deviation	Range
/i/	Heed	243.511	42.513	200.998 - 286.024
/a/	Hod	260.933	44.192	216.741 - 305.125
/u/	Who'd	236.711	47.268	189.433 - 283.979
/o/	Hode	265.978	49.678	216.300 - 315.656
/e/	Hayd	267.244	41.214	226.030 - 308.458
/ɛ/	Head	195.556	34.216	161.350 - 229.782
/ɪ/	Hid	192.889	36.881	156.008 - 229.770
/æ/	Had	271.622	37.188	234.434 - 308.810

Stimuli development

The study used the 12 nonword stimuli developed for Coalson and Byrd (2015) and maintained the same carefully controlled characteristics. That is, each nonword bore minimal resemblance to actual English words, while still retaining common properties of Standard American English. The following properties were controlled for to the greatest extent possible: (a) neighborhood density and frequency (e.g., Anderson, 2007; Vitevitch, Luce, Pisoni, & Auer, 1999), (b) phonotactic probability (e.g., Anderson & Byrd, 2008; Vitevitch & Sommers, 2003), (c) word shape and phonetic complexity (e.g., Howell, Au-Yeung, Yaruss, & Eldridge, 2006; cf. Coalson, Byrd, & Davis, 2012), (d) syllable frequency (e.g., Levelt & Wheeldon, 1994; Cholin, Dell, & Levelt, 2011), (e) uniqueness point (e.g., Ozdemir, Roelofs, & Levelt, 2006), and (f) clarity of syllable boundary in medial consonant clusters (see Wheeldon & Morgan, 2002, pp. 516-517). Prior to testing, Coalson and Byrd (2015) presented all 12 nonword stimuli aurally to 10 AWNS who did not participate in the experimental portion of the study. A 5-point Likert scale described by Gathercole (1995: 1 = *very unlike a real word*, 5 = *very like a real word*) was used to assess word-likeness. All stimuli received mean scores between *unlike a real word* to *neutral*.

Target phonemes

The same six consonants [*l, m, ʃ, v, z, f*] used in Coalson and Byrd (2015) were used in the present study. Unlike Coalson and Byrd (2015), eight vowels [*i, a, u, o, e, ɛ, ɪ, æ*] were also represented in the nonword stimuli, detailed below in Table 2. These 14 sounds were presented within the nonword stimuli across 12 total blocks, with each block presented in randomized order. Each block was comprised of 25 trials, all of which occurred in a fixed, random order within the block. Ten of the 25 trials included the target CVCCVC nonword (allowing for six true-positive tokens, four false-positive tokens). The six true positive tokens included all of the sounds within the CVCCVC nonword. The remaining fifteen trials within the block included five presentations of each of the three foils in fixed random order, with 1 to 2 false negative responses per monosyllabic foil. Therefore, across all 12 blocks, each participant had the opportunity to provide 72 true positive responses for the target nonword (6 segmental positions x 12 nonwords).

Table 2.Lexical and phonological properties of target stimuli with associated foils (Coalson & Byrd, 2015).

Block	Experiment 1	Experiment 2	Foil 1	Foil 2	Foil 3
1	/'viʃ.fuz/	/viʃ.'fuz/	/ʃɛv/	/zom/	/laf/
2	/'zæl.ʃov/	/zæl.'ʃov/	/vif/	/miʃ/	/ləz/
3	/'ʃiv.lom/	/ʃiv.'lom/	/vuz/	/fəʃ/	/mɛl/
4	/'fæz.mul/	/fæz.'mul/	/vim/	/zof/	/ʃəl/
5	/'lam.vef/	/lam.'vef/	/fɛʃ/	/miv/	/zol/
6	/'muf.zoʃ/	/muf.'zoʃ/	/faz/	/vim/	/ʃəl/
7	/'foʃ.vul/	/foʃ.'vul/	/ʃaz/	/zɪf/	/miv/
8	/'lev.mof/	/lev.'mof/	/vəl/	/faʃ/	/zim/
9	/'mæz.fuv/	/mæz.'fuv/	/vɛf/	/ʃom/	/zel/
10	/'ʃɛm.liz/	/ʃɛm.'liz/	/fuʃ/	/zev/	/mæl/
11	/'vul.ziʃ/	/vul.'ziʃ/	/ʃaf/	/fɛv/	/lom/
12	/'zɪf.ʃom/	/zɪf.'ʃom/	/vul/	/feʃ/	/məz/

Training task

Participants completed training tasks prior to the experimental portion to mitigate practice effects (see Smits-Bandstra, 2010). In addition, due to the novelty of the stimuli, sufficient practice ensured greater accuracy during the nonword repetition task (see Coalson & Byrd, 2015; see Gupta, 2003; *n*=18). Finally, practice training prior to the experimental portion allowed participants to identify and generate target nonwords with greater accuracy during the experimental portion based upon the on-screen location of the non-orthographic cue with which they were trained to associate that specific stimulus (Coalson & Byrd, 2015, see Figure 1). This also served to ensure that participants were not visually scanning for the stimulus or recruiting visuo-spatial aspects of short-term memory.

Silent phoneme monitoring task

As in Coalson and Byrd (2015), the experimental portion of the present study used silent monitoring of phonemes within stimuli to assess between-groups latency and accuracy of responses. The silent monitoring task for each phoneme took place in a block immediately following its associated training task. The following instructions were presented on the screen: "You will hear a sound, then you will see one of the words, or the speaker icon, appear in its corner. Do not say the word aloud. If the sound is in the word, press 'Yes'. If not, press 'No'. After you press 'Yes' or 'No,' you will be cued to say the word aloud." The sequence for silent phoneme monitoring tasks reflected that of Coalson and Byrd (2015), and is detailed in Figure 2.

Analyses

As in Coalson and Byrd (2015), responses were omitted prior to analysis if: (a) response time fell greater than 2 standard deviations below that participant's talker group mean for that particular stimulus, (b) no response was provided, (c) the participant simultaneously verbalized the response while providing a manual response, or (d) the post-trial verbal response was characterized by phonemic or stress errors. Inter-rater reliability between and within participants was ensured for all phonological subtests using Likert ratings via inter-class coefficient.

Investigators assessed reaction time latencies by conducting a 2x6 repeated measures mixed-model, with talker group (i.e., AWS or AWNS) as the between-group factor, and response times related to phoneme position (i.e., C_1-V_1-C_2-C_3-V_2-C_4) as the within-group factor (Coalson & Byrd, 2015). Similar to Coalson and Byrd (2015), investigators conducted planned pairwise comparisons to examine mean latencies of consonants and vowels which occur prior to the syllable boundary (i.e., C_1, V_1, C_2) in the presence of initial stress, both within and between talker groups.

As in Coalson and Byrd (2015), investigators examined false negative errors using 2x6 repeated measures mixed-model ANOVA, with talker group as the between-group factor and phoneme position error rates as the within-group factor. Investigators conducted two separate one-way ANOVAs for phonemic and stress errors to examine post-trial verbal error data between groups (Coalson & Byrd, 2015). Talker group served as the independent variable, and error rate served as the dependent variable. Individual alpha levels were set at .05 for all ANOVA and pairwise comparisons. Investigators also conducted independent *t*-tests to compare group performance on participant age, and phonological processing subtest performance. All planned and post-hoc comparisons used Fisher's Least Significant Difference adjusted *p*-values, and Greenhouse-Geisser adjusted *F*-values were reported for all ANOVAs that did not meet assumptions of sphericity.

Training Task							
Nonword Stimuli	Auditory Stimuli	Auditory Presentation	Visual Stimuli	Visual Cue	Participant Response	Order	No. of Exposures
1. Repetition Phase							
Target	Heard target	Simultaneous with Visual Cue	Orthographic, Corner A	Corner A	Repeat aloud	Random	4
Foil 1	Heard foil	Simultaneous with Visual Cue	Orthographic, Corner B	Corner B	Repeat aloud	Random	4
Foil 2	Heard foil	Simultaneous with Visual Cue	Orthographic, Corner C	Corner C	Repeat aloud	Random	4
Foil 3	Heard foil	Simultaneous with Visual Cue	Orthographic, Corner D	Corner D	Repeat aloud	Random	4
2. Identification Phase							
Target	Heard target	Simultaneous with Visual Cue	Neutral Icon, Corner A	Corners A-D	Point to corner	Random	4
Foil 1	Heard foil	Simultaneous with Visual Cue	Orthographic, Corner B	Corners A-D	Point to corner	Random	4
Foil 2	Heard foil	Simultaneous with Visual Cue	Orthographic, Corner C	Corners A-D	Point to corner	Random	4
Foil 3	Heard foil	Simultaneous with Visual Cue	Orthographic, Corner D	Corners A-D	Point to corner	Random	4
3. Generation Phase							
Target	None	None	Neutral Icon, Corner A	Corner A	Say nonword	Random	4
Foil 1	None	None	Orthographic, Corner B	Corner B	Say nonword	Random	4
Foil 2	None	None	Orthographic, Corner C	Corner C	Say nonword	Random	4
Foil 3	None	None	Orthographic, Corner D	Corner D	Say nonword	Random	4
4. [Repeat Steps 1-3 with 2 exposures to each nonword per phase]							
5. [If participant failed to accurately produce word 2/2 times during final generation phase, repeat step 4]							
6. [If participant failed to accurately produce word 2/2 times during final generation phase, repeat step 5]							

Silent Phoneme Monitoring Task							
Nonword Stimuli	Auditory Stimuli	Audio Presentation	Visual Stimuli	Visual Cue	Participant Response*	Order	No. of Trials
Target	[f, z, m, l, v, ʃ, i, a, u, o, e, ɛ, ɪ, or æ]	900 ms Before Visual Cue	Neutral Icon, Corner A	Corner A	1. Press Yes/No 2. Say word	Fixed Random	10
Foil 1	Varied	900 ms Before Visual Cue	Orthographic, Corner B	Corner B	1. Press Yes/No 2. Say word	Fixed Random	5*
Foil 2	Varied	900 ms Before Visual Cue	Orthographic, Corner C	Corner C	1. Press Yes/No 2. Say word	Fixed Random	5*
Foil 3	Varied	900 ms Before Visual Cue	Orthographic, Corner D	Corner D	1. Press Yes/No 2. Say word	Fixed Random	5*

Figure 1. Structure of a training task and silent phoneme monitoring task within a single block (adapted from Coason & Byrd, 2015).

Note: Three-phase training procedures detailed in Coalson and Byrd (2015) are similar to paradigms described by Levelt and colleagues (Cholin, Levelt, & Schiller, 2006; Cholin et al., 2011; Levelt & Wheeldon, 1994; and Gupta (2003). **Number of Yes/No responses for foils varied across block*

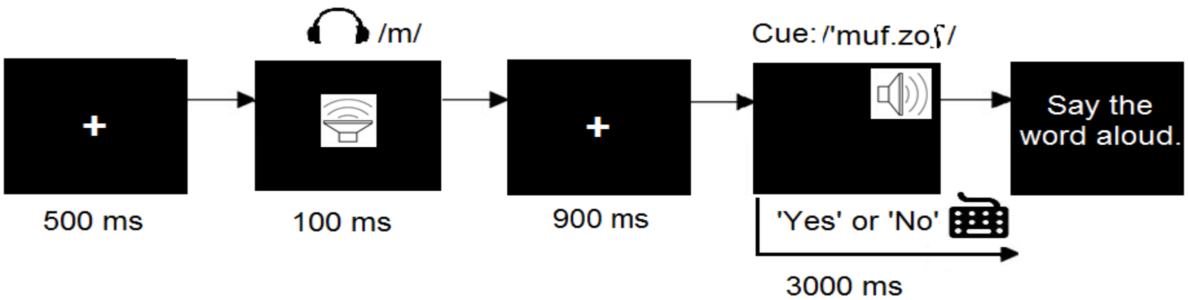

Figure 2. Silent phoneme monitoring task for nonword stimuli (Coalson & Byrd, 2015).

Experiment 2: Initial Vowel in the Absence of Syllabic Stress

In accordance with Research Question 2, the purpose of Experiment 2 was to examine whether AWS differ temporally from AWNS when encoding the medial vowel (V_1) with non-initial stress in CVCCVC stimuli. Participant intake procedures, criteria for inclusion or exclusion, experimental paradigm, and analyses for Experiment 2 were identical to those of Experiment 1 (also see Coalson & Byrd, 2015). Experiment 2 differed from Experiment 1 in only one aspect – assignment of stress to the second syllable of target nonwords, rather than the first syllable. This modification allowed investigators to assess the speed and accuracy of encoding in the absence of initial stress, both between and within groups. Participants included in Experiment 2 had not participated in the first experiment and were only presented with nonwords with non-initial stress to diminish the difficulties inherent to learning and monitoring identical phonemic sequences with both initial stress, then non-initial stress, or vice versa.

CHAPTER III: RESULTS

Experiment 1

Eight participants (AWNS, *n* = 4; AWS, *n* = 4) completed Experiment 1, which examined silent phoneme monitoring abilities of bisyllabic nonword stimuli with initial stress. Each group included three males and one female with no significant difference in age (AWNS, *M* = 26.00, *SD* = 6.78; AWS, *M* = 32.00, *SD* = 18.07). No significant differences were observed between AWNS and AWS during pre-experimental testing (see Table 3; *p*-value range = .448 to .838).

Table 3.
Mean age and performance on phonological processing subtests for talker groups.

	Experiment 1		Experiment 2	
	AWNS	AWS	AWNS	AWS
N [m, f]	4 [3, 1]	4 [3, 1]	4 [2, 2]	4 [3, 1]
Age	26.00 (6.78)	32.00 (18.07)	21.75 (0.96)	25.50 (6.45)
Nonword Repetition	19 (1.77)	20.50 (0.50)	17.00 (1.47)	17.00 (1.63)
Real Word Segmentation	14.50 (3.23)	13.75 (1.38)	12.00 (3.34)	14.75 (2.93)
Nonword Segmentation	26.75 (1.25)	25.25 (1.37)	24.50 (2.02)	26.75 (2.01)
Forward Digit Span	21.25 (1.11)	20.50 (0.87)	21.25 (1.65)	17.25 (1.11)
Backward Digit Span	18.50 (0.87)	17.00 (1.91)	16.00 (1.08)	14.00 (1.91)

Note: Mean age calculated in years and standard deviation (in parentheses). Nonword Segmentation, Nonword Segmentation, and Forward Digit Span from *Comprehensive Test of Phonological Processes – Second Edition* (CTOPP-2; Wagner, Torgeson, Rashotte, & Pearson, 2013) and reported as mean raw scores and standard error (in parentheses). Backward Digit Span derived from modified instructions from Forward Digit Span subtest. Real Word Segmentation from *Comprehensive Test of Phonological Processes* (CTOPP; Wagner, Torgeson, & Rashotte, 1999) and reported as mean raw scores and standard error (in parentheses).

To reduce the likelihood that reaction time was influenced by inaccurate retrieval, processing, or identification of phonemes within the target nonword, reaction time data were based on accurate nonverbal monitoring as well as phonetically accurate, fluent post-trial verbal responses. Individual tokens were excluded from reaction time analysis if they met the following criteria:

(a) False Negative: Participant manually responded ‘no’ when target phoneme was present
(b) Phonemic Error: Post-trial verbal response included one or more phonemic errors, delayed verbal response (2+ seconds post-trial response), overlapping verbal response with button press, and/or no verbal or manual response
(c) Stress Error: Post-trial response was produced with inaccurate syllabic stress
(d) Disfluent Response: Post-trial verbal response contained a stuttering-like disfluency (i.e., sound-syllable repetition, audible sound prolongation, and/or inaudible sound prolongation) or a typical disfluency (i.e., interjection, revision; Yairi & Ambrose, 2005)

(e) Technical Error: Response was unintelligible or coding was not possible due to audio/video difficulties, non-speech event (e.g., yawn, cough), and/or software error[1].

From the collected tokens (*N* = 576; AWNS, *n* = 288, AWS, *n* = 288), 50 tokens (8.7%) were removed based on false negative responses (AWNS, *n* = 15; AWS, *n* = 35); 48 tokens (8.7%) were excluded based on phonemic error (AWNS, *n* = 18; AWS, *n* = 30), 23 tokens were (4.0%) excluded based on stress error (AWNS, *n* = 20; AWS, *n* = 3), 3 tokens were (0.5%) excluded based on disfluency during post-trial response (AWNS, *n* = 1; AWS, *n* = 2), and 69 tokens (12.0%) were excluded due to technical error (AWNS, *n* = 36; AWS, *n* = 33). The final data corpus included 383 usable tokens (AWNS, *n* = 198; AWS, *n* = 185).

Reaction time latencies

As depicted in Figure 3, a mixed-model repeated measures ANOVA was conducted to assess the reaction time latencies of AWNS and AWS during silent phoneme monitoring of C_1, V_1, and C_2 phonemes within the first syllable of bisyllabic nonwords with initial stress. There was no significant group by position interaction, Wilks' Lambda = .928, $F(2, 5) = 5.16$, $p = 0.170$, $\eta\rho^2 = 0.928$, and no significant main effect for talker group ($F < 1$). However, there was a significant main effect for consonant position, Wilks' Lambda = .928, $F(2, 5) = 21.09$, $p = 0.046$, $\eta\rho^2 = 0.981$.

Planned comparisons indicated that mean latencies between talker groups did not differ for C_1, V_1, or C_2 phonemes. However, planned within-group comparisons revealed significantly longer latencies for V_1 relative to C_1 for both AWNS (C_1:*M* = 859.22 ms, *SE* = 180.45 ms; V_1: *M* = 992.71, *SE* = 153.57) and AWS (C_1: *M* = 941.10 ms, *SE* = 180.45; V_1: *M* = 1209.87 ms, *SE* = 153.57). Neither group exhibited a significant difference between V_1 and C_2 phonemes, with the C_2 monitored with relatively longer latencies (AWNS: *M* = 1038.22 ms, *SE* = 221.49; AWS: *M* = 1298.56 ms, *SE* = 221.49) compared to the medial, stress-bearing vowel.

False negative errors

A mixed-model repeated measure ANOVA was conducted to assess the accuracy of nonverbal identification of phonemes for AWNS and AWS during the silent phoneme monitoring task. As depicted in Figure 3, no significant main effect was found for phoneme position ($F < 1$), and no significant group by position interaction was detected, Wilks' Lambda = .948, $F(2, 5) = 7.290$, $p = 0.125$, $\eta\rho^2 = 0.948$. However, a significant main effect for group was found Wilks' Lambda = .928, $F(1, 6) = 7.978$, $p = 0.030$, $\eta\rho^2 = 0.571$, with AWS exhibiting

[1]During data collection, experimenters discovered that presentation of the /m/ phoneme was incorrect during the silent phoneme monitoring task. Tokens associated with these manual and verbal responses were considered a technical error and therefore excluded prior to final analyses.

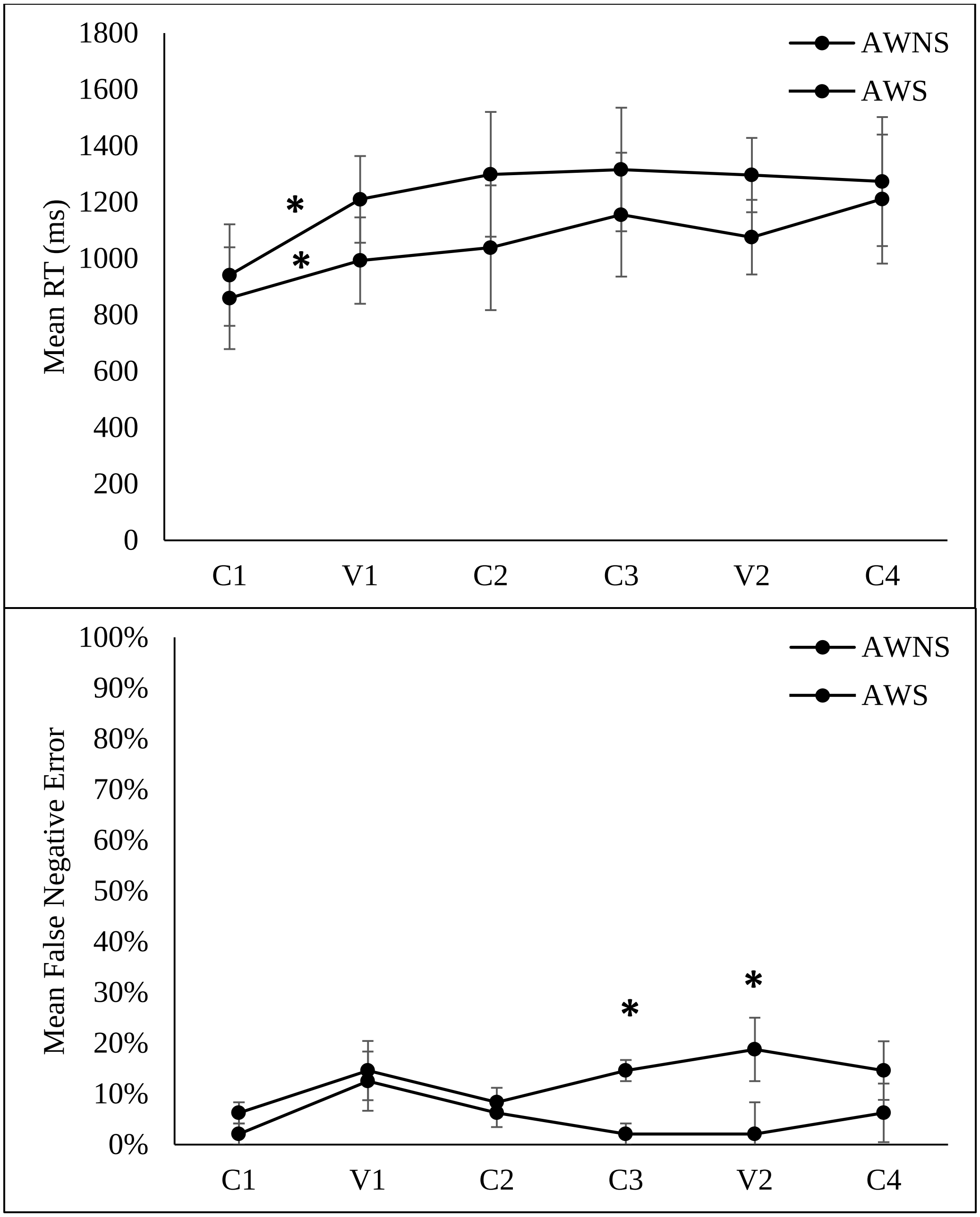

Figure 3. Mean reaction time (RT) latencies and mean false negative errors for adults who do and do not stutter (AWS, AWNS) when silently monitoring phonemes in bisyllabic nonwords with initial stress.

significantly higher false negative errors (*M* = 12.8%, *SE* =1.9%) than AWNS (*M* = 5.2%, *SE* = 1.9%). Planned comparisons revealed no significant differences in C_1, V_1, or C_2 phonemes between AWNS (C_1:*M* = 2.1%, *SE* = 2.1%; V_1: *M* = 12.5%, *SE* = 5.8%; C_2: *M* = 6.3%, *SE* = 2.8%) and AWS (C_1:*M* = 6.3%, *SE* = 2.1%; V_1: *M* = 14.6%, *SE* = 8.3%; C_2: *M* = 8.3%, *SE* = 2.8%). Significant differences were revealed for onset of the second, unstressed syllable (C_3), with AWS (*M* = 14.6%, *SE* = 2.1%) exhibiting greater errors on average than AWNS (*M* = 2.1%, *SE* = 2.1%).

Experiment 2

Eight participants (AWNS, n = 4, AWS, n = 4) not included in Experiment 1 completed Experiment 2, which examined silent phoneme monitoring abilities of bisyllabic nonword stimuli with non-initial stress. AWNS included two males and two females and the AWS included three males and one female, with no significant difference in age (AWNS, M = 21.75, SD = 0.96; AWS, M = 25.50, SD = 6.45). No significant differences were observed between AWNS and AWS during pre-experimental testing (see Table 3; p-value range = .091 to .99).

As in Experiment 1, to reduce the likelihood that reaction time was influenced by inaccurate retrieval, processing, or identification of phonemes within the target nonword, reaction time data was based on accurate nonverbal monitoring as well as phonetically accurate, fluent post-trial verbal responses. From the collected tokens (N = 576; AWNS, n = 288, AWS, n = 288), 80 tokens (14.9%) were removed based on false negative responses (AWNS, n = 37; AWS, n = 43); 100 tokens (17.4%) were excluded based on phonemic error (AWNS, n = 49; AWS, n = 51), 0 tokens were excluded based on stress error (AWNS, n = 0; AWS, n = 0), 27 tokens (4.7%) were excluded based on disfluency during post-trial production (AWNS, n = 6; AWS, n = 21), and 60 tokens (10.4%) were excluded due to technical error (AWNS, n = 36; AWS, n = 24). The final data corpus for reaction time analysis included 309 usable tokens (AWNS, n = 160; AWS, n = 149).

Reaction time latencies

As depicted in Figure 3, a mixed-model repeated measures ANOVA was conducted to assess the reaction time latencies of AWNS and AWS during silent phoneme monitoring of C_1, V_1, and C_2 phonemes within the first syllable of bisyllabic nonwords with initial stress. There was no significant group by position interaction, ($F < 1$), and no significant main effect for group ($F < 1$), or consonant position, Wilks' Lambda = .904, $F(2, 5) = 3.75$, $p = 0.224$, $\eta\rho^2 = 0.904$.

False negative errors

A mixed-model repeated measure ANOVA was conducted to assess the accuracy of nonverbal identification of phonemes for AWNS and AWS during the silent phoneme monitoring task. As depicted in Figure 3, no significant main effect was found for phoneme position ($F < 1$), or talker group, ($F < 1$). No significant group by position interaction was detected, Wilks' Lambda = .323, $F(2, 2.559) = 2.867$, $p = 0.077$, $\eta\rho^2 = 0.323$.

Post-trial production errors

Phonemic and stress-based errors produced for AWNS and AWS after each silent phoneme monitoring trial were examined using two separate one-way ANOVAs for each experiment (Experiment 1: initial stress; Experiment 2: non-initial stress), with talker group (AWNS, AWS) as the independent variable and error rate as the dependent variable. As depicted in Figures 4 and 5, no significant differences were revealed for phonemic errors, F<1, or stress-based errors, $F(3, 15) = 1.766$, $p = 0.207$, $\eta^2 = 0.295$ between AWNS and AWS.

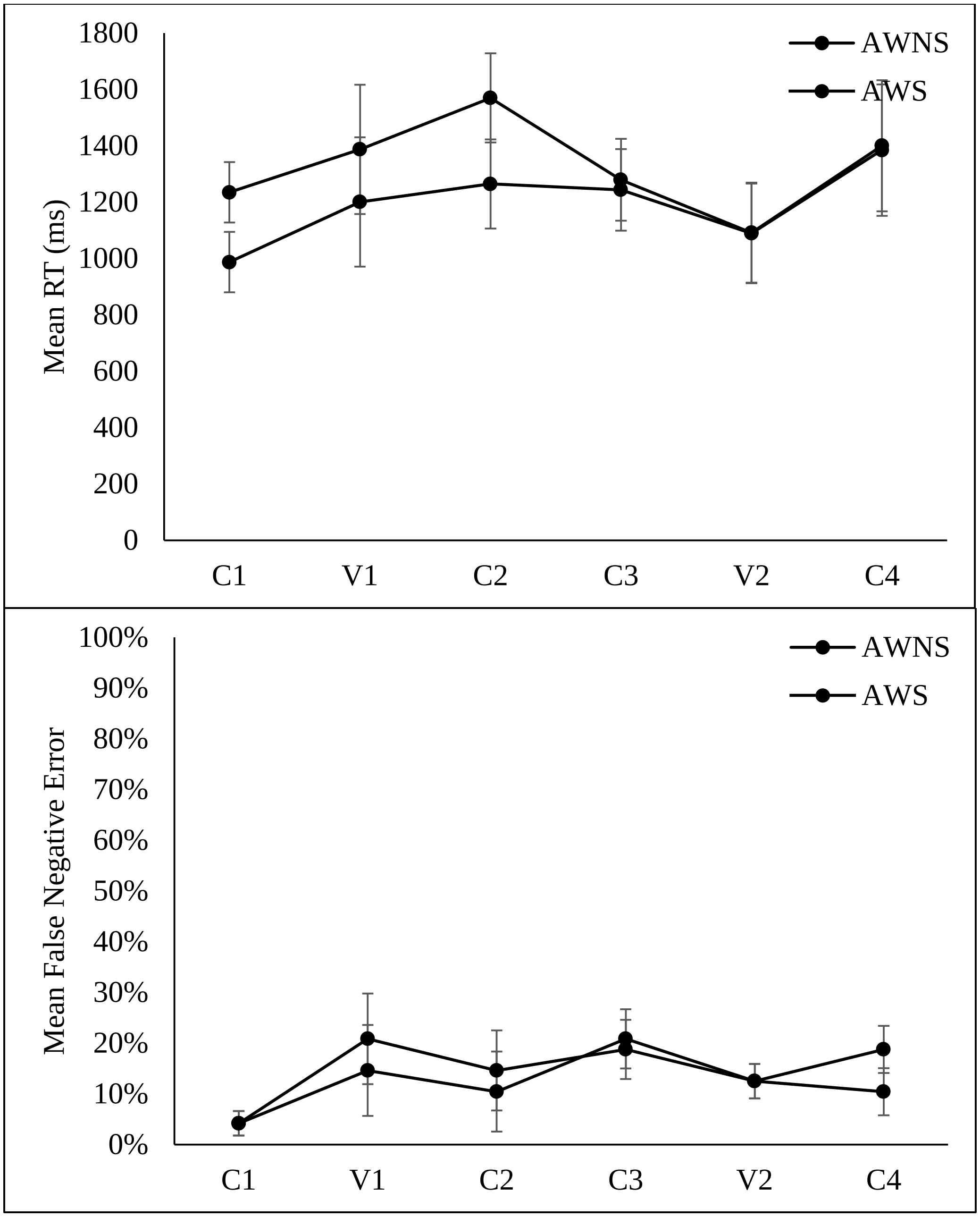

Figure 4. Mean reaction time (RT) latencies and mean false negative errors for adults who do and do not stutter (AWS, AWNS) when silently monitoring phonemes in bisyllabic nonwords with non-initial stress.

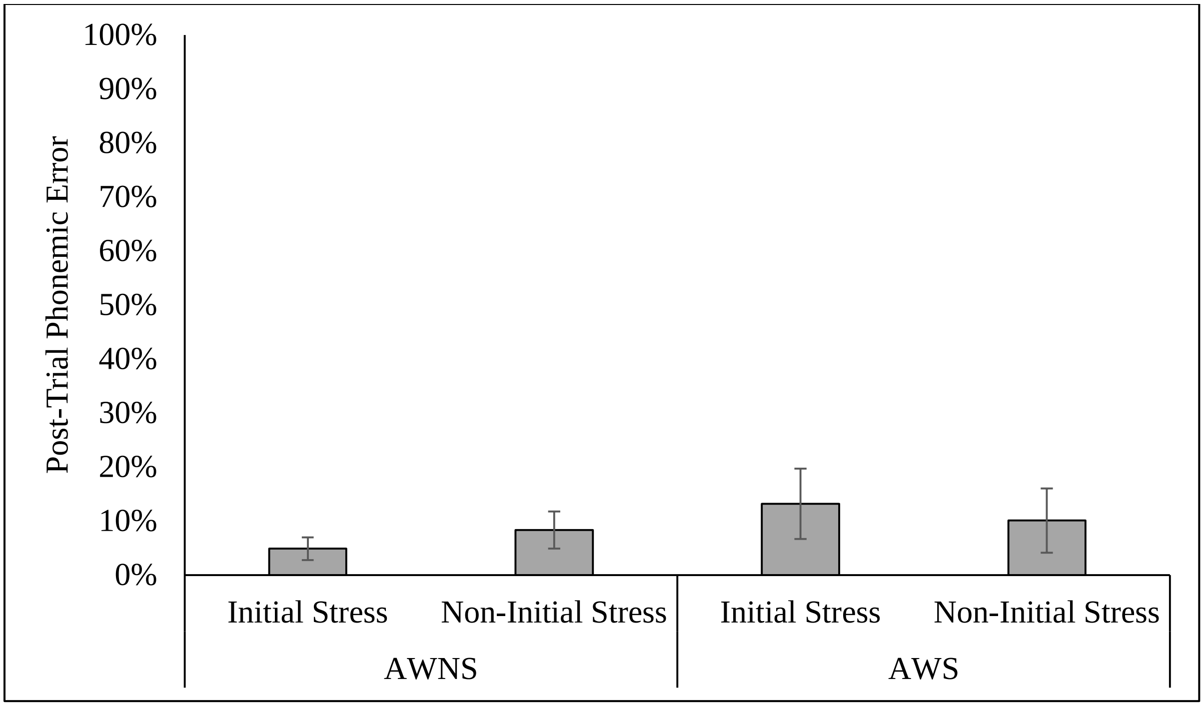

Figure 5. Mean phonemic errors produced by adults who do and do not stutter (AWS, AWNS) during post-trial verbal responses.

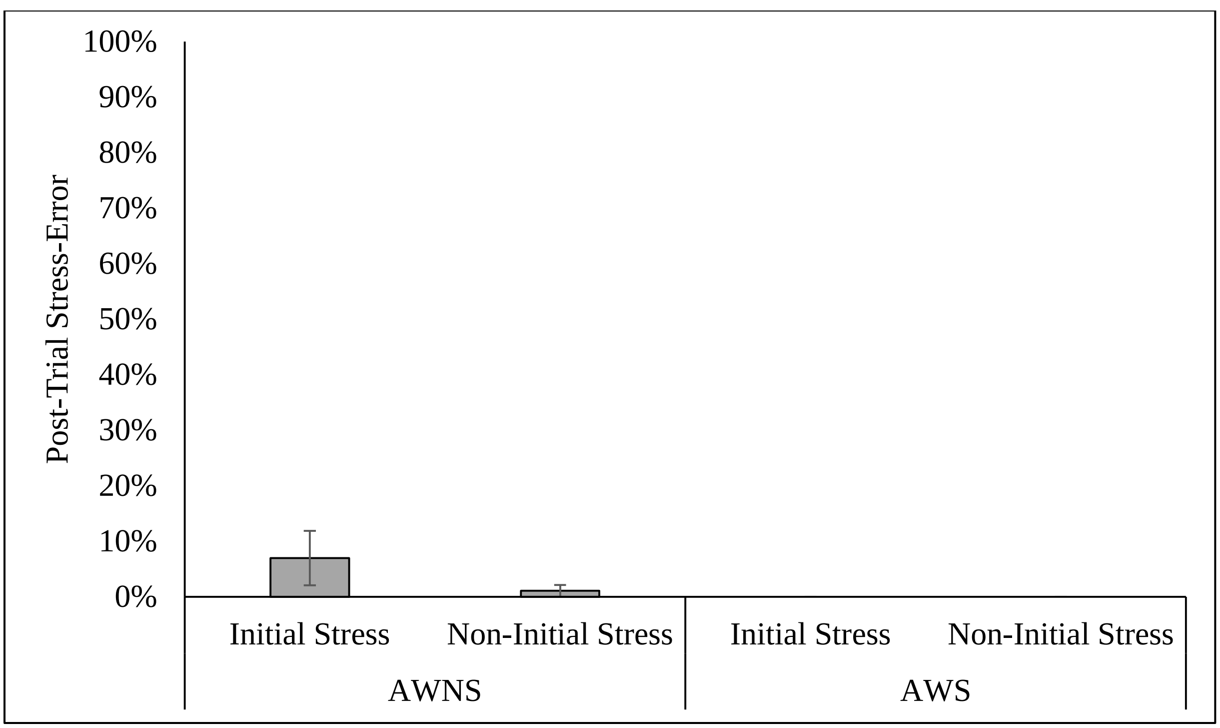

Figure 6. Mean stress-assignment errors produced by adults who do and do not stutter (AWS, AWNS) during post-trial verbal responses.

CHAPTER IV: DISCUSSION

A growing collection of literature suggests stuttering may be related to difficulties encoding the segmental properties of speech (for review, see Byrd, Wolk, & Davis, 2007) and, to some extent, the metrical properties of speech (for further information see Coalson & Byrd, 2015). One theoretical model of stuttering - Wingate's (1988) Fault-Line Hypothesis - predicts that the combined influence of segmental and metrical encoding delays assignment of the stress-bearing medial vowel, and serves as the critical point of breakdown in AWS during speech planning and production. However, prior research has not directly examined the time course of phonological encoding of the initial medial vowel (V_1) in individuals who stutter. Therefore, the purpose of this study was to examine how specific metrical properties of words such as stress assignment may impact the speed of phonological encoding of target phonemes, specifically stressed and unstressed vowels, in AWS and AWNS. The following research questions were posed:

1) Do AWS differ from AWNS during phonological encoding of initial medial vowels (V_1) in CVCCVC stimuli with initial stress?
2) Do AWS differ from AWNS during phonological encoding of initial medial vowels (V_1) in CVCCVC stimuli with non-initial stress?

Experiment 1 examined whether AWS differ from AWNS in encoding of initial medial vowels (V_1) with stimuli characterized by initial stress. Results from Experiment 1 indicated that phonological encoding of sounds within the first, stressed syllable of CVCCVC stimuli occurred with similar latencies between talker groups. Both talker groups exhibited significant delays between C_1 and V_1, indicating a relative delay encoding the medial vowel (V_1), with no significant difference between V_1 and C_2 for either group. In Experiment 2, stress was removed from the initial syllable. Upon removal of initial stress, AWS and AWNS did not significantly differ in speed when monitoring the phonemes within the unstressed syllable. Across both experiments, false negative error rates did not differ between or within groups for any phoneme position, and post-trial verbal production errors did not differ between groups. Together, the lack of observed group differences in either study suggests that the influence of stress was similar regardless of talker group classification.

In Experiment 1, overall monitoring latencies for phoneme positions which have been included in extant literature (C_1, C_2, C_3, and C_4) mirrored the time course of speech planning reported for AWNS (e.g., Wheeldon & Levelt, 1995; Wheeldon & Morgan, 2002) and AWS (Coalson & Byrd, 2015; Sasisekaran et al., 2006; Sasisekaran et al., 2013). That is, the four consonants within C_1-V_1-C_2-C_3-V_2-C_4 structure were monitored at increasingly slower latencies throughout the word, suggesting left-to-right assembly of the preverbal speech plan, with significant delays between C_1 and C_2. However, it should be noted that the bisyllabic stimuli lists in prior studies were comprised exclusively or predominately of initial stress, and therefore the unique contribution of the medial vowel remained speculative prior to the current study.

Examination of the intervening vowel provided evidence that the significant C_1 - C_2 delay reported in in previous studies may reflect slower encoding of the stress-bearing vowel (V_1) relative to the initial consonant (C_1), rather than latency differences following medial vowel and final consonant of the syllable (C_2). This contrasts with data reported by Wheeldon and Morgan (2002) which indicated a significant latency difference between C_1 and V_1, and well as V_1 and C_2. Findings are consistent with error pattern data reported by Treiman and Danis (1988), who suggested that syllables are coded in terms of onset (C_1) and rime (V_1+C_2), with the onset less vulnerable to error. Despite the disparate significance levels between Wheeldon and Morgan and the present study, the magnitude of monitoring latencies between phonemes remains similar. Mean V_1-C_2 latency differences found in the present study for AWNS (45.51 ms) and AWS (88.69 ms) were similar to those reported by Wheeldon and Morgan (77 ms). C_1-V_1 differences in AWNS in Experiment 1 (133.49 ms) were also comparable to AWNS in Wheeldon and Morgan (124 ms). In AWS, however, reaction times in the present study nearly doubled the latency differences reported by Wheeldon and Morgan (268.77 ms). It is possible that within-group significance, and perhaps between group differences, would have been achieved with sufficient power and reduction of standard error to resemble the significant latency differences reported by Wheeldon and Morgan between each of the first three phonemes within initial, stressed syllables. Nonetheless, preliminary findings in Experiment 1 of the present study are largely consistent with previous studies of AWNS for phonemes within initial, stressed syllables, and the overall time course from onset to coda is consistent with previous reports in both AWNS and AWS. The lack of between group differences in Experiment 1, and Experiment 2 is also congruent with data reported by Burger and Wijnen (1999), in that the absence of between group differences when monitoring the stressed vowel does not support Wingate's (1988) contention that AWS exhibit significantly greater difficulty encoding initial, stress-bearing vowels.

Experiment 2 examined the monitoring latencies of the phonemes within the first syllable of stimuli without initial stress for AWNS and AWS. In contrast to stimuli with initial stress in Experiment 1, neither group - AWNS or AWS - monitored C_1, V_1 or C_2 with significantly different speed. These data suggest that in the absence of stress, initial segments are assigned to the syllabic frame with no significant delay between phonemes. Minimal data prior to this study is available to predict the monitoring latency of phonemes within initial syllables but without initial stress. Schiller (2005) reported no C_1-C_2 latency differences in the absence of initial stress for AWNS. In contrast, Coalson and Byrd (2015) reported that both AWNS and AWS exhibit significant C_1-C_2 latencies in the presence *and* absence of initial stress, with no group differences observed. Mean latency differences between the present study and Coalson and Byrd indicate a similarity in the overall time course of encoding for initial, unstressed syllables, but differences in magnitude of latencies between phonemes. For instance, Coalson and Byrd reported C_1-C_2 latency differences of approximately 160 ms for AWS and AWNS (without examination of the medial vowel). AWS and AWNS in Experiment 2 of the present study exhibited nearly a twofold increase in C_1-C_2 latencies (AWNS: 277.61 ms; AWS: 335.06 ms). Again, sufficient power may have been necessary to reduce the overall standard error and to achieve statistical significance between phoneme positions and between groups. In addition, AWNS in Experiment 2 showed the greatest delay between C_1 and V_1 (214.01 ms; V_1-C_2 = 63.60 ms) within the initial unstressed syllable, similar to words with initial stress in Experiment 1, while AWS demonstrated comparable latency differences between C_1-V_1 (152.39 ms) and V_1-C_2 (182.67 ms). Thus, while statistical significance was not achieved, further investigation may reveal greater temporal 'separation' of phonemes within unstressed initial syllables during speech planning in AWS.

Four main methodological limitations in the present study warrant cautious interpretation of results across both experiments. First and foremost, with data from a total of 16 participants (eight participants per experiment), investigators were unable to achieve sufficient power in order to demonstrate the predicted significance between AWS and AWNS in Experiment 1. Large effect sizes for non-significant differences in monitoring latencies between groups ($\eta\rho^2$ =

0.928) and between phonemes ($\eta\rho^2 = 0.981$) in Experiment 1, and group by phoneme interaction in Experiment 2 ($\eta\rho^2 = 0.904$) provide some support that this may be the case (Cohen, 1988). Second, and unlike prior experiments (i.e., Experiment 3 of Wheeldon & Morgan, 2002), the present study required participants to identify isolated sounds via auditory cues. The purpose of this cueing technique was to eliminate the potential for visual representation of the target phoneme and/or nonword target by the participant, as well as the influence of orthographic transparency on response time. Auditory cueing has been used successfully with consonant targets in Coalson and Byrd (2015). Nonetheless, this approach may have posed a unique challenge for participants when required to discriminate between vowel targets during training and experimental tasks. Third, the present study had an insufficient number of participants across groups to account for the influence of potential covariates during post-trial analyses (see Stevens, 1996). This may be a particularly relevant consideration with regard to error analyses the AWS. For example, Coalson and Byrd found performance on digit span tasks correlated with false negative responses and post-trial phonemic responses for stimuli with non-initial stress for AWS (but not AWNS). In Experiment 2, differences between AWNS and AWS during forward digit span tasks approached significance ($p = .091$), suggesting that similar considerations may be necessary to adequately assess differences during experimental tasks with greater statistical power. Finally, data attrition was a significant factor. Technical errors limited the number of usable tokens, and the number of post-trial production errors was substantially higher than previous studies. As suggested by Cutler, van Ooiijen, Norris, and Sanchez-Casas (1996), phoneme monitoring tasks with vowel targets typically yield higher error rates and are perceived by participants to be more difficult, and response criterion may require adjustments to account for these perceived difficulties. Future studies should consider these limitations with the use of an increased number of participants, and perhaps explore alternate stimulus presentation methods such as orthographic and/or visual cues.

Conclusion

The purpose of this study was to examine differences in reaction time latencies between AWS and AWNS when monitoring initial, medial vowels which do and do not carry metrical stress. Previous models of stuttering implicate stress-bearing medial vowels as problematic for AWS and prone to disrupt the time course of speech planning and, ultimately, fluent speech production. Findings from the present study suggest that in both the presence and absence of syllabic stress, AWS monitored medial vowels with speed and accuracy commensurate with AWNS. Despite methodological limitations, these preliminary data provide minimal support for the notion that simultaneous encoding of segmental and metrical information imposes greater demand during speech planning for individuals who stutter.

REFERENCES

Ambrose, N. G., Yairi, E., Loucks, T. M., Seery, C. H., & Throneburg, R. (2015). Relation of motor, linguistic and temperament factors in epidemiologic subtypes of persistent and recovered stuttering: Initial findings. *Journal of Fluency Disorders, 45*, 12-26.

Anderson, J. D. (2007). Phonological neighborhood and word frequency effects in the stuttered disfluencies of children who stutter. *Journal of Speech, Language & Hearing Research*, *50*(1), 229-247.

Anderson, J., & Byrd, C. (2008). Phonotactic probability effects in children who stutter. *Journal of Speech, Language & Hearing Research, 51*(4), 851-866.

Arndt, J., & Healey, E. C. (2001). Concomitant disorders in school-age children who stutter. *Language, Speech, and Hearing Services in Schools, 32*(2), 68-78.

Blood, G. W., Ridenour, V. J., Qualls, C. D., & Hammer, C. S. (2003). Co-occurring disorders in children who stutter. *Journal of Communication Disorders, 36*(6), 427-448.

Bloodstein, O., & Bernstein Ratner, N.B. (2008). *A handbook on stuttering – Sixth edition*. Clifton Park, NY: Thomson Delmar Learning.

Burger, R., & Wijnen, F. (1999). Phonological encoding and word stress in stuttering and nonstuttering subjects. *Journal of Fluency Disorders, 24*(2), 91-106.

Byrd, C. T., Conture, E. G., & Ohde, R. N. (2007). Phonological priming in young children who stutter: Holistic versus incremental processing. *American Journal of Speech-Language Pathology*, *16*(1), 43-53.

Byrd, C. T., Wolk, L., & Davis, B. L. (2007). Role of phonology in childhood stuttering and its treatment. In E. Conture & E. Curlee (Eds.), *Stuttering and related disorders of fluency – Third Edition* (pp.168-182). New York, NY: Thieme.

Cholin, J., Dell, G. S., & Levelt, W. M. (2011). Planning and articulation in incremental word production: syllable-frequency effects in english. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(1), 109-122.

Cholin, J., Levelt, W. J., & Schiller, N. O. (2006). Effects of syllable frequency in speech production. *Cognition*, *99,* 205-235.

Coalson, G. A., & Byrd, C. T. (2015). Metrical encoding in adults who stutter. *Journal of Speech, Language, and Hearing Research, 58,* 601-621.

Coalson, G. A., Byrd, C. T., & Davis, B. L. (2012). The influence of phonetic complexity on stuttered speech. *Clinical Linguistics & Phonetics*, *26*(7), 646-659.

Coalson, G. A., Byrd, C. T., & Rives, E. (in press). Academic, clinical, and educational experiences of self-identified fluency specialists. *Perspectives on Fluency and Fluency Disorders*.

Cohen, J. W. (1988). *Statistical power analysis for the behavioral sciences – Second edition*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Cooper, E. B., & Cooper, C. S. (1985). Clinician attitudes toward stuttering: A decade of change (1973-1983). *Journal of Fluency Disorders*, *10*(1), 19-33.

Cooper, E. B., & Cooper, C. S. (1996). Clinician attitudes towards stuttering: Two decades of change. *Journal of Fluency Disorders, 21*(2), 119-135.

Craig, A., Blumgart, E., & Tran, Y. (2009). The impact of stuttering on the quality of life in adults who stutter. *Journal of Fluency Disorders*, *34*(2), 61-71.

Cutler, A., van Ooijen, B., Norris, D., & Sanchez-Casas, R. (1996). Speeded detection of vowels: A cross linguistic study. *Perception and Psychophysics, 58,* 807-822.

Gabel, R. M. (2014). School speech-language pathologists' experiences with stuttering: An Ohio survey. *eHearsay, 3*, 5-29.

Gathercole, S. E. (1995). Is nonword repetition a test of phonological memory or long-term knowledge? It all depends on the nonwords. *Memory & Cognition*, *23*(1), 83-94.

Gupta, P. (2003). Examining the relationship between word learning, nonword repetition, and immediate serial recall in adults. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 56,* 1213-1236.

Hartsuiker, R. J. (2007). Studies of verbal self-monitoring: The perceptual loop model and beyond. In A. Meyer, L. Wheeldon, & A. Knott (Eds.), *Automaticity and control in language processing* (pp. 93-121). Psychology Press: New York, NY.

Hillenbrand, J., Getty, L., Clark, M., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *The Journal of the Acoustical Society of America*, *97*(5), 3099-3111.

Howell, P., Au-Yeung, J., Yaruss, J., & Eldridge, K. (2006). Phonetic difficulty and stuttering in English. *Clinical Linguistics & Phonetics*, *20*(9), 703-716.

Indefrey, P., & Levelt, W. J. (2004). The spatial and temporal signatures of word production components. *Cognition*, *92*(1), 101-144.

Jansma, B. M., & Schiller, N. O. (2004). Monitoring syllable boundaries during speech production. *Brain and Language, 90*(1), 311-317.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*, 1-75.

Levelt, W. J. M., & Wheeldon, L. (1994). Do speakers have access to a mental syllabary? *Cognition*, *50,* 239-269.

Natke, U., Grosser, J., Sandrieser, P., & Kalveram, K. T. (2002). The duration component of the stress effect in stuttering. *Journal of Fluency Disorders*, *27*(4), 305-318.

Natke, U., Sandrieser, P., van Ark, M., Pietrowsky, R., & Kalveram, K. T. (2004). Linguistic stress, within-word position, and grammatical class in relation to early childhood stuttering. *Journal of Fluency Disorders*, *29*(2), 109-122.

Nippold, M. A. (2002). Stuttering and phonology: Is there an interaction?. *American Journal of Speech-Language Pathology*, *11*(2), 99.

O'Brian, S., Packman, A., Onslow, M., & O'Brian, N. (2004). Measurement of stuttering in adults: Comparison of stuttering-rate and severity-scaling methods. *Journal of Speech, Language, and Hearing Research, 47,* 1081–1087.

Ozdemir, R., Roelofs, A., & Levelt, W. J. M. (2006). Perceptual uniqueness point effects in monitoring internal speech. *Cognition, 105*, 457-465.

Postma, A., & Kolk, H. (1993). The Covert Repair Hypothesis: Prearticulatory repair processes in normal and stuttered disfluencies. *Journal of Speech and Hearing Research*, *36,* 472-487.

Pirello, K., Blumstein, S. E., & Kurowski, K. (1997). The characteristics of voicing in syllable-initial fricatives in American English. *The Journal of the Acoustical Society of America, 101*(6), 3754-3765.

Sasisekaran, J., Brady, A., & Stein, J. (2013). A preliminary investigation of phonological encoding skills in children who stutter. *Journal of Fluency Disorders*, *38*(1), 45-58.

Sasisekaran, J., de Nil, L. F., Smyth, R., & Johnson, C. (2006). Phonological encoding in the silent speech of persons who stutter. *Journal of Fluency Disorders, 31*, 1-21.

Schiller, N. O. (2005). Verbal self-monitoring. In A. Cutler (Ed.), *Twenty-first century psycholinguistics: Four cornerstones* (pp. 245-261). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Smits-Bandstra, S. (2010). Methodological considerations in the measurement of reaction time in persons who stutter. *Journal of Fluency Disorders*, *35*(1), 19-32.

Spencer, C., & Weber-Fox, C. (2014). Preschool speech articulation and nonword repetition abilities may help predict eventual recovery or persistence of stuttering. *Journal of Fluency Disorders, 41*, 32-46.

St. Louis, K. O., & Durrenberger, C. H. (1993). What communication disorders do experienced clinicians prefer to manage? *ASHA*, *35*(12), 23.

Stevens, J. (1996). *Applied multivariate statistics for the social sciences – Third edition.* Mahwah, NJ: Lawrence Erlbaum.

Treiman, R., & Danis, C. (1988). Short-term memory errors for spoken syllables are affected by the linguistic structure of the syllables. *Journal of Experimental Psychology*, *14*, 145-152.

Vitevitch, M. S., Luce, P. A., Pisoni, D. B., & Auer, E. T. (1999). Phonotactics, neighborhood activation, and lexical access for spoken words. *Brain and Language, 68,* 306-311.

Vitevitch, M. S., & Sommers, M. S. (2003). The facilitative influence of phonological similarity and neighborhood frequency in speech production in younger and older adults. *Memory and Cognition, 31*, 491-504.

Wagner, R. K., Torgeson, J. K., & Rashotte, C. A. (1999). *Comprehensive Test of Phonological Processes (CTOPP).* Austin, TX: Pro-Ed.

Wagner, R. K., Torgeson, J. K., Rashotte, C. A., & Pearson, N. A. (2013). *Comprehensive Test of Phonological Processes (CTOPP-2).* Austin, TX: Pro-Ed.

Wheeldon, L. R., & Levelt, W. J. (1995). Monitoring the time course of phonological encoding. *Journal of Memory and Language, 34*(3), 311-334.

Wheeldon. L. R., & Morgan, J. L. (2002). Phoneme monitoring in internal and external speech. *Language and Cognitive Processes, 17*, 503-535.

Wijnen, F., & Boers, I. (1994). Phonological priming effects in stutterers. *Journal of Fluency Disorders*, *19*(1), 1-20.

Wingate, M. E. (1988). *The structure of stuttering: A psycholinguistic analysis,* Springer-Verglag, New York, NY.

Wolk, L., Edwards, M. L., & Conture, E. G. (1993). Coexistence of stuttering and disordered phonology in young children. *Journal of Speech, Language, and Hearing Research, 36*(5), 906-917.

Yairi, E., & Ambrose, N. G. (2005). *Early childhood stuttering for clinicians by clinicians*. Austin, TX: Pro-Ed.

Yairi, E., & Ambrose, N. (2013). Epidemiology of stuttering: 21st century advances. *Journal of Fluency Disorders*, *38*(2), 66-87.

VITA

Allison E. Jacobs, a native of Boulder, Colorado, received her bachelor's degree in Speech Language and Hearing Sciences in 2013. As her interest in the field grew, she made the decision to pursue a master's degree in Communication Disorders at the Louisiana State University. She anticipates earning her master's degree in May 2016 and thereafter plans to earn clinical certification as a speech-language pathologist.