JSLHR

Research Article

Metrical Encoding in Adults Who Do and Do Not Stutter

Geoffrey A. Coalson[a] and Courtney T. Byrd[b]

Purpose: The purpose of this study was to explore metrical aspects of phonological encoding (i.e., stress and syllable boundary assignment) in adults who do and do not stutter (AWS and AWNS, respectively).
Method: Participants monitored nonwords for target sounds during silent phoneme monitoring tasks across two distinct experiments. For Experiment 1, the 22 participants (11 AWNS, 11 AWS) silently monitored target phonemes in nonwords with initial stress. For Experiment 2, an additional cohort of 22 participants (11 AWNS, 11 AWS) silently monitored phonemes in nonwords with noninitial stress.
Results: In Experiment 1, AWNS and AWS silently monitored target phonemes in initial stress stimuli with similar speed and accuracy. In Experiment 2, AWS demonstrated a within-group effect that was not present for AWNS. They required additional time when monitoring phonemes immediately following syllable boundary assignment in stimuli with noninitial stress. There was also a between-groups effect, with AWS exhibiting significantly greater errors identifying phonemes in nonwords with noninitial stress than AWNS.
Conclusions: Findings suggest metrical properties may affect the time course of phonological encoding in AWS in a manner distinct from AWNS. Specifically, in the absence of initial stress, metrical encoding of the syllable boundary may delay speech planning in AWS and contribute to breakdowns in fluent speech production.

Multiple theories of stuttering suggest phonological encoding may compromise the fluency of speech production (e.g., Howell, 2011; Postma & Kolk, 1993; Wingate, 1988). *Phonological encoding*, as described by Levelt and colleagues (Levelt, 1989; Levelt, Roelofs, & Meyer, 1999), refers to the assembly of abstract speech plans prior to speech production. Within this theoretical model of speech production, two types of information are activated simultaneously to generate the preverbal speech plan: segmental information (i.e., the sounds within a word) and metrical information (i.e., syllable stress and syllable boundary assignment). Upon activation, rapid reintegration of segmental and metrical properties—or syllabification—is required to facilitate fluent speech production. Delayed or inefficient assignment of segmental and/or metrical properties during the relatively brief time frame allowed for syllabification (e.g., Indefrey & Levelt, 2004) would postpone the timely preparation of the speech plan and potentially disrupt the fluency of spoken output.

[a]Louisiana State University, Baton Rouge
[b]The University of Texas at Austin
Correspondence to Geoffrey A. Coalson: gcoals1@lsu.edu

Editor: Jody Kreiman
Associate Editor: Hans-Georg Bosshardt
Received April 22, 2014
Revision received October 21, 2014
Accepted January 21, 2015
DOI: 10.1044/2015_JSLHR-S-14-0111

There are both experimental and descriptive data that suggest segmental information may be underspecified in children and adults who stutter (AWS; e.g., Anderson, 2007; Anderson & Byrd, 2008; Bosshardt, 1993; Byrd, Conture, & Ohde, 2007; Sasisekaran & Byrd, 2013; Sasisekaran, de Nil, Smyth, & Johnson, 2006; cf. Hennessey, Nang, & Beilby, 2008; Nippold, 2012; Vincent, Grela, & Gilbert, 2012). By comparison, the relationship between stuttering and metrical aspects of phonological encoding has largely been restricted to descriptive data. Within these limited data, a significant relationship has been observed with respect to the metrical property of stress. Stuttering typically occurs on initial, stressed syllables (e.g., Natke, Grosser, Sandrieser, & Kalveram, 2002; Natke, Sandrieser, van Ark, Pietrowsky, & Kalveram, 2004; Prins, Hubbard, & Krause, 1991; Wingate, 1988). These descriptive findings suggest that both segmental and metrical aspects of phonological encoding may contribute to the difficulties persons who stutter have establishing and/or maintaining fluent speech. Review of data from typically fluent speakers lends further support to the investigation of segmental and metrical properties in individuals who stutter at the level of phonological encoding because each holds the potential to alter the time course of syllabification prior to the production of speech.

Disclosure: The authors have declared that no competing interests existed at the time of publication.

Segmental and Metrical Encoding in Nonstuttering Adults

A fundamental component of Levelt et al.'s (1999) model of speech production is the ability to monitor phonological code generated during syllabification prior to production. As noted by Levelt et al., the use of a silent phoneme monitoring task allows the integrity of the phonological code to be examined via the internal self-monitoring system. During silent phoneme monitoring tasks, the participant identifies via nonverbal response (e.g., button press) the presence or absence of a target phoneme within a target word. For example, Wheeldon and colleagues (Wheeldon & Levelt, 1995; Wheeldon & Morgan, 2002) presented typically fluent adults with semantically related stimuli to cue participants' silent generation of target words. All target words shared a similar, bisyllabic word shape (i.e., $C_1VC_2C_3VC_4$, where C = consonant, V = vowel) and carried either initial stress (e.g., *magnet*: C_1 = /m/, C_2 = /g/, C_3 = /n/, C_4 = /t/) or noninitial stress (e.g., *canteen*: C_1 = /k/, C_2 = /n/, C_3 = /t/, C_4 = /n/). For both studies, each consonant was identified at incrementally slower rates (i.e., $C_1 < C_2 < C_3 < C_4$), indicating that segmental information is available to the speaker in a left-to-right manner during the phonological encoding process.

In addition to incremental assembly of segmental information, significant latency differences were observed in both studies (Wheeldon & Levelt, 1995; Wheeldon & Morgan, 2002) when participants monitored phonemes that flanked either initial stress assignment (i.e., C_1 and C_2: 55 ms and 109 ms, respectively) or syllable boundary assignment (i.e., C_2 and C_3: 56 ms and 63 ms, respectively). These delays were interpreted as a reflection of the additional time required to process each metrical property of speech during phonological encoding. However, response time data in each study were averaged across stimuli with mixed stress assignment (i.e., Wheeldon & Levelt: first syllable = 15; second syllable = 5, as reported by Kruyskamp, 1961; Wheeldon & Morgan: first syllable = 12; second syllable = 6, as reported by Wells, 2008), with no significant differences observed within the final syllable (i.e., C_3–C_4). Thus, it is difficult to determine from these data whether stress assignment contributes equally to segmental assignment across all positions within a word or whether the observed effects of syllable boundary assignment occur independently of metrical stress configuration.

To isolate the influence of the metrical properties during phonological encoding, Schiller (2005) manipulated stress assignment of bisyllabic $C_1VC_2C_3VC_4$ target words during a silent phoneme monitoring task in Dutch-speaking adults. Findings indicated that when internally generated words involved initial stress, a significant latency difference was found between C_1 and C_2, with no appreciable difference between C_2, C_3, and C_4. In contrast, for words with noninitial stress, there was no discernible difference in monitoring latencies between C_1 and C_2. Rather, the only significant latency difference found was for phonemes flanking the syllable boundary, with C_3 taking significantly longer to monitor than C_2. Schiller interpreted these data in the context of the rightward segmental monitoring reported in studies by Wheeldon and colleagues (Wheeldon & Levelt, 1995; Wheeldon & Morgan, 2002). Specifically, he stated that if speakers assemble and monitor phonological speech code from left to right, the first-encountered metrical property may require additional time to compute, be it initial stress (i.e., $C_1 < C_2$) or in the absence of initial stress, the first syllable boundary (i.e., $C_2 < C_3$). Schiller further explained that during the time required to formulate the first-encountered metrical property, all subsequent segmental and metrical properties are afforded sufficient time to become fully activated and, thus, are immediately available after the initial metrical property has been completed.

Segmental and Metrical Encoding in Individuals Who Stutter

Schiller's (2005) explanation of phoneme monitoring differences specific to metrical properties may account for observed patterns of stuttered speech. Descriptive data provide a link between moments of stuttering and metrical stress, with the observation that stuttering occurs most frequently on initial, stressed syllables of utterances (e.g., Natke et al., 2004; Prins et al., 1991; Wingate, 1988). However, more data are needed to determine whether the metrical properties of initial stress as well as syllable boundary uniquely contribute to overt moments of stuttering. If delayed encoding of metrical properties compromise fluent speech, individuals who stutter may exhibit distinct phoneme monitoring patterns when assigning either initial stress or the initial syllable boundary.

Experimental data to support the independent contribution of each metrical property to phonological encoding are limited because silent phoneme monitoring studies examining both C_1–C_2 (i.e., initial stress) and C_2–C_3 (i.e., syllable boundary) differences in individuals who stutter have been restricted to stimuli comprising only initial stress. For example, Sasisekaran et al. (2006) used a silent phoneme monitoring task with English-speaking AWS and adults who do not stutter (AWNS; n = 10 and 11, respectively) using $C_1VC_2C_3VC_4$ targets with initial stress. Participants heard a verbal prompt (e.g., "Please respond to the /sə/ sound in the following picture") and were then presented with one of 14 pictured stimuli in random order. Overall, AWS were significantly slower to monitor phonemes than AWNS. Using a similar paradigm, Sasisekaran, Brady, and Stein (2013) studied children 10 to 14 years of age and found that those who stutter (n = 9) also demonstrated significantly slower mean reaction times than fluent peers (n = 9) when silently monitoring phonemes in words with initial stress. The significantly slower reaction times reported across studies lend support to delayed segmental encoding in individuals who stutter. However, the observed between-groups differences in reaction time should be interpreted with caution, given two critical methodological factors.

First, Smits-Bandstra (2010) noted a tendency for participants who stutter to exhibit delayed reaction time on the basis of a simple manual response (i.e., button press) during early trials, with differences diminishing with increased practice. According to her report (p. 23, Table 2), between-groups differences would be expected across a variety of response types (e.g., button press, production of CV targets, arm movement) when provided a restricted number of trials ($n = 16$) and a limited number of pretrial exposures ($n = 3$ to 5 practice trials). Second, as acknowledged by Sasisekaran et al. (2006, p. 15), use of real words as stimuli cannot completely discount potential group differences in lexical access (e.g., word frequency: Anderson, 2007; Newman & Ratner, 2007; semantic information: Bosshardt & Fransen, 1996). For example, Newman and Ratner (2007, p. 205) reported that the magnitude of the response time difference was greater in AWS than AWNS when producing words of low versus high frequency. Thus, to better isolate phonological encoding from potential underlying motoric variability or differences in lexical retrieval, the present study provided increased practice prior to experimental tasks and used nonword stimuli rather than real-word stimuli.

Beyond the between-groups differences observed in Sasisekaran and colleagues (2006, 2013), intriguing within-group trends suggest that the potential influence of metrical properties during phonological encoding may be unique to the stuttering cohorts. For example, C_1–C_2 reaction time differences in Sasisekaran et al. (2006, p. 12) indicate that AWS required more time to process segments separated by the initial, stress-bearing vowel (C_1–C_2 latency difference: 306.5 ms) relative to typically fluent adults (C_1–C_2 latency difference: 214 ms). These data lend support to the notion that segmental encoding may be delayed in AWS due to an apparent difficulty in the assignment of metrical stress. These data are also compatible with previous descriptions of increased stuttering on the initial stressed syllable of words (e.g., Natke et al., 2004; Wingate, 1988). With respect to syllable boundary, Sasisekaran et al. (2013, p. 55) reported that children who stutter required significantly longer to monitor phonemes across the syllable boundary (C_2–C_3 latency difference: $M = 170$ ms) compared with typically fluent peers (C_2–C_3 latency difference: $M = -47$ ms) even in the presence of significantly protracted C_1–C_2 latencies. These results suggest that despite the presence of initial stress and significant C_1–C_2 latency differences, additional time may be required for children who stutter to monitor phonemes following the syllable boundary. Together, these findings suggest the metrical properties of speech planning may be uniquely challenging for persons who stutter.

Thus, the purpose of the present study was to examine the metrical properties of stress and syllable boundary assignment in the phonological encoding of AWS and AWNS. The limited experimental data available suggest that assignment of metrical properties (i.e., initial stress: Sasisekaran et al., 2006; initial syllable boundary: Sasisekaran et al., 2013) may delay encoding in individuals who stutter in a manner that is distinct from typically fluent peers. Unlike prior studies in which data were aggregated across initial stress stimuli (Sasisekaran et al., 2006, 2013), in the present study we isolated the independent influence of stress and syllable boundary assignment similar to Schiller (2005) by positioning each as the first-encountered metrical property in two separate experiments. In Experiment 1, AWS and AWNS monitored phonemes in bisyllabic stimuli in which stress was assigned to the first syllable. In Experiment 2, a distinct group of participants monitored phonemes within the same bisyllabic stimuli in the absence of initial stress, exposing the syllable boundary as the first-encountered metrical property.

The present study also differs from previous studies in two additional, critical ways. First, to control the influence of practice effects (i.e., Smits-Bandstra, 2010), increased training was provided prior to participation in the experimental silent phoneme monitoring tasks. Second, the potential contribution of lexical access (Sasisekaran et al., 2006) on performance was attenuated by having participants monitor nonword rather than real-word stimuli. We predicted that with sufficient training, as well as limited lexical influence, overall between-groups reaction time differences would be minimal or absent. However, the proposed difficulties with metrical encoding in AWS suggest that phoneme monitoring patterns within groups should be dissimilar. Specifically, we predicted AWS would exhibit disproportionately slower monitoring of phonemes that flank either initial stress (C_1–C_2, Experiment 1) and/or syllable boundary (C_2–C_3, Experiment 2).

Method

Experiment 1: Influence of Initial Stress

Participants

The present study was approved by the authors' university institutional review board. Informed consent was obtained for each participant and financial compensation was provided. Participants completed two 90-minute sessions. Data collected during the first 90-minute session determined inclusion in the experimental portion of the study, which was completed during the second 90-minute session. During the first session the following tasks were completed: (a) self-report questionnaire, (b) hearing, vision, and handedness screening, (c) phonological processing subtests, (d) baseline reaction time task, (e) nonlinguistic auditory monitoring task, (f) identification of target phonemes in isolation, and (g) talker group classification tasks.

Self-report questionnaire. The self-report questionnaire provided the following information for each participant: age, gender, race and/or ethnicity, primary language(s) spoken, socioeconomic status (as determined by level of education and job status), prior or current medical difficulties, history of reading and/or learning difficulties, current use of antipsychotic medications, and a summary of speech and/or language treatment history.

Hearing, vision, and handedness screening. Hearing, vision, and handedness were assessed to ensure that manual responses to the auditory and/or orthographic stimuli during experimental tasks (i.e., baseline reaction time, nonlinguistic auditory monitoring, training, and silent phoneme monitoring tasks) were not influenced by peripheral factors unrelated to the processes of interest. All participants completed binaural pure-tone hearing screening (American Speech-Language-Hearing Association [ASHA], 1997), a vision screening (U.S. Department of Health and Human Services, 1996), and a handedness inventory (Oldfield, 1971). If a participant demonstrated left-hand dominance on handedness screening (i.e., < −40), yes and no button assignment was reversed prior to all manual response tasks to confirm that reaction times were not influenced by use of the nondominant hand.

Phonological processing subtests. A series of pre-experimental criterion-referenced tasks were administered to ensure that the participants within or between groups did not present with significant deficits in phonological encoding and/or working memory abilities. Nonword repetition and onset identification of real words and nonwords were measured using subtests from the Comprehensive Test of Phonological Processing (Wagner, Torgesen, & Rashotte, 1999). Participants were also asked to provide another word that starts with the same first sound after each test item to assess word generation on the basis of initial phoneme. In addition, participants completed two criterion-referenced tasks to assess basic rhyme identification and generation abilities (Sasisekaran & Byrd, 2013: 10 real words, five nonwords). Participants were also required to complete forward and backward digit span tasks (Wechsler Adult Intelligence Scale–Third Edition; Wechsler, 1997).

Baseline reaction time task. To establish baseline reaction time, participants responded to a visual icon presented in one of four corners of the computer monitor at randomized intervals (i.e., 500, 1,000, 1,500, and 2,000 ms) by pressing a button as quickly as possible.

Nonlinguistic auditory monitoring task. The purpose of the auditory monitoring task was to assess the perceptual monitoring abilities of participants to confirm that reaction-time latencies reflected processes related to phonological encoding rather than to underlying differences in auditory discrimination (see Demonet et al., 1992). Similar to Sasisekaran et al. (2006), each participant heard 96 auditory stimuli consisting of a four-tone sequence. Half of the stimuli (n = 48) were composed of four identical 5 kHz tones, each tone 100 ms in duration with 50 ms between tones. The remaining half (n = 48) included three 5-kHz tones and one 1-kHz tone. After brief familiarization with the low and high tones, participants were instructed to press the yes button as quickly as possible if they heard the high tone anywhere in the series and to press the no button as quickly as possible if they did not hear the high tone. Each trial consisted of an orienting cross in the center of the screen (500 ms), followed by simultaneous auditory presentation of the tone series accompanied with a visual cue in one of four corners of the computer monitor. All 96 stimuli were presented in a random order (i.e., 48 yes and 48 no responses, 12 per auditory position).

Identification of target phonemes in isolation. In addition to pure-tone hearing screening, participants were required to accurately identify the six target phonemes included in the experimental tasks in isolation, similar to their presentation during the subsequent silent phoneme monitoring task. This identification task was included to further confirm accuracy of auditory discrimination for the target phonemes. Participants first listened to each target phoneme in a randomized order with simultaneous visual presentation of the corresponding grapheme (e.g., heard /m/ and saw the letter M) with no required response (6 phonemes × 4 presentation, n = 24 total exposures). After the initial 24 exposures, participants heard each target phoneme without visual input and were instructed to verbally identify the corresponding letter (i.e., heard /m/, responded m). Participants were required to identify all 24 phonemes in isolation with 100% accuracy in order to be included for participation in the experimental tasks.

Talker group classification. Classification as an AWS required (a) self-identification as an individual who stutters with onset reported prior to the age of 7 years, and (b) prior diagnosis of stuttering by a licensed speech-language pathologist. If the participant had not received a prior formal diagnosis of stuttering, in addition to the participant's self-identification as a person who stutters, AWS status was further confirmed by the first author, an ASHA-certified, licensed speech-language pathologist.

Severity ratings were completed on a conversational and also a reading sample using a 9-point severity scale (1 = *no stuttering*, 2 = *very mild stuttering*, 5 = *moderate stuttering*, 9 = *extremely severe stuttering*; O'Brian, Packman, Onslow, & O'Brian, 2004). Neutral topics were included in conversational samples (e.g., movies, books, holiday plans) and reading samples (e.g., "The Rainbow Passage," "The Grandfather Passage"). Each AWS scored 2 or higher on either the conversational or reading sample. All AWS self-identified as individuals who stutter and six of 11 had previously received a diagnosis of stuttering. All AWNS participants received a severity rating of no higher than 1 (i.e., *no stuttering*) for both samples, and they also did not self-identify as an AWS or report previous diagnoses of stuttering (see Table 1). Interrater reliability of stuttering severity for speech samples was determined by the first author and an undergraduate research assistant trained in disfluency-count analysis. Six of the 22 participants (27%; three AWNS, three AWS) were selected at random. For AWS, interrater reliability for both speech samples were within 1 point for the three participants, with all participants receiving a score of 2 or higher (intraclass coefficient [ICC] = .95 for conversation, .97 for reading); intrarater reliability ratings were within 1 point (ICC = .96 for conversation; .97 for reading). There was 100% agreement for the severity ratings for AWNS during conversation and reading samples across all three participants, all of whom received a score of 1 (i.e., *no stuttering*).

Inclusionary and exclusionary criteria. To qualify for inclusion, participants had to meet the following criteria: (a) 18 years of age or older, (b) no history of cognitive, developmental, or neurological disturbances, (c) no reported current use of antipsychotic medications, (d) no reported or observed speech or language concerns with the exception of stuttering for the AWS participants, and (e) reported as well as demonstrated nativelike proficiency in English.

Table 1. Participant characteristics for adults who do and do not stutter (AWS, AWNS, respectively) in Experiments 1 and 2.

	Conversational sample		Reading sample					Conversational sample		Reading sample			
	Severity score	Stuttering severity	Severity score	Stuttering severity	Previous Dx	Self-ID		Severity score	Stuttering severity	Severity score	Stuttering severity	Previous Dx	Self-ID
Experiment 1							Experiment 2						
AWS-1	5	mod	5	mod	N	Y	AWS-12	5	mod	3	mild	Y	Y
AWS-2	4	mild–mod	7	sev	Y	Y	AWS-13	2	very mild	2	very mild	Y	Y
AWS-3	4	mild–mod	7	sev	Y	Y	AWS-14	4	mild–mod	2	very mild	N	Y
AWS-4	6	mod-sev	7	sev	Y	Y	AWS-15	3	mild	3	mild	Y	Y
AWS-5	2	very mild	2	very mild	N	Y	AWS-16	2	very mild	2	very mild	N	Y
AWS-6	3	mild	3	mild	N	Y	AWS-17	1	none	2	very mild	Y	Y
AWS-7	4	mild–mod	3	mild	Y	Y	AWS-18	1	none	2	very mild	N	Y
AWS-8	3	mild	1	none	N	Y	AWS-19	1	none	2	very mild	Y	Y
AWS-9	8	very sev	7	sev	Y	Y	AWS-20	9	ex sev	9	ex sev	Y	Y
AWS-10	3	mild	2	very mild	N	Y	AWS-21	3	mild	2	very mild	Y	Y
AWS-11	1	none	2	very mild	Y	Y	AWS-22	5	mod	5	mod	Y	Y
AWNS-1	1	none	1	none	N	N	AWNS-12	1	none	1	none	N	N
AWNS-2	1	none	1	none	N	N	AWNS-13	1	none	1	none	N	N
AWNS-3	1	none	1	none	N	N	AWNS-14	1	none	1	none	N	N
AWNS-4	1	none	1	none	N	N	AWNS-15	1	none	1	none	N	N
AWNS-5	1	none	1	none	N	N	AWNS-16	1	none	1	none	N	N
AWNS-6	1	none	1	none	N	N	AWNS-17	1	none	1	none	N	N
AWNS-7	1	none	1	none	N	N	AWNS-18	1	none	1	none	N	N
AWNS-8	1	none	1	none	N	N	AWNS-19	1	none	1	none	N	N
AWNS-9	1	none	1	none	N	N	AWNS-20	1	none	1	none	N	N
AWNS-10	1	none	1	none	N	N	AWNS-21	1	none	1	none	N	N
AWNS-11	1	none	1	none	N	N	AWNS-22	1	none	1	none	N	N

Note. Severity score indicates score on 9-point scale of stuttering severity (O'Brian et al., 2004). Previous Dx = previous diagnosis of stuttering; Self-ID = participant self-identification as an adult who stutters; AWS = adults who stutter; AWNS = adults who do not stutter; mod = moderate; mild–mod = mild to moderate; mod sev = moderately severe; sev = severe; ex sev = extremely severe.

Participants were excluded if they (a) did not pass hearing or vision screenings, (b) identified target phonemes in isolation with less than 100% accuracy, or (c) had a mean baseline manual reaction time that was beyond two standard deviations of the mean reaction time of the participant's talker group (*Z* scores). The two talker groups were also balanced for age and gender. Appendix A depicts the number of participants included and excluded from initial recruitment for each talker group in Experiment 1 and Experiment 2.

The final participant pool for Experiment 1 included 22 adults (11 AWS, seven men, four women, age range: 18 to 41 years, $M = 22$ years; 11 AWNS, seven men, four women, age range: 18 to 26 years, $M = 20$ years) with nativelike proficiency in Standard American English (see Appendix A).

Stimuli Development

Nonword development. Twelve nonword stimuli were constructed with minimal resemblance to real words while retaining phonotactically legal segmental properties. Phonological composition of nonwords stimuli were identical across Experiments 1 and 2 except for the assignment of syllabic stress (Experiment 1: initial stress; Experiment 2: noninitial stress). Target phonemes within nonwords (i.e., C_1, C_2, C_3, and C_4) were selected on the basis of four exclusionary criteria. First, target phonemes were not included if the phoneme did not occur in all positions (i.e., onset, medial, and coda) in the English language (e.g., [h, j, ŋ]). Second, target phonemes were not included if aural identification in isolation was difficult without inclusion of additional acoustic properties that also cue word position (e.g., aspiration during voiceless stops). Third, consonants that were difficult to discriminate in certain positions without an accompanying vowel (i.e., voiced stops, rhotics) were not included to prevent coarticulatory input that may cue position. Finally, phonemes with alternative grapheme representations were avoided due to the inhibitory effect of orthographic redundancy on reaction time during phoneme monitoring tasks (Damian & Bowers, 2003; Dijkstra, Frauenfelder, & Schreuder, 1993). Of the remaining consonants that did not meet exclusionary criteria, six target phonemes were selected: [l, m, ʃ, v, z, f]. First-syllable vowels were balanced for duration by assigning six tense and six lax vowels. All second-syllable vowels were tense to maintain further balance. Each of the six target phonemes occupied each of the four consonant positions twice and occurred eight times across the 12 nonword stimuli (see Table 2).

Eight additional factors were considered during construction of nonword stimuli when balancing stimuli for lexical, phonological, and phonetic properties of speech with known or potential influence on speed, accuracy, or fluency of response in stuttering and nonstuttering adults, including neighborhood density and frequency (e.g., Anderson, 2007; Vitevitch, Luce, Pisoni, & Auer, 1999), phonotactic probability (e.g., Anderson & Byrd, 2008; Vitevitch & Sommers, 2003), word shape and phonetic complexity (e.g., Howell, Au-Yeung, Yaruss, & Eldridge, 2006; cf. Coalson, Byrd, & Davis, 2012), syllable frequency (e.g., Cholin, Dell, & Levelt, 2011; Levelt & Wheeldon, 1994), orthographic transparency (e.g., Damian & Bowers, 2003; Dijkstra et al., 1993), uniqueness point (e.g., Ozdemir, Roelofs, & Levelt, 2006), and syllable boundary clarity of medial consonant clusters (see Wheeldon & Morgan, 2002, pp. 516–517). Word-likeness for these 12 nonwords was then assessed by presenting each auditorally to 10 typically fluent adults who were asked to apply a 5-point Likert scale described in Gathercole (1995): 1 = *very unlike a real word*, 5 = *very like a real word*. Mean word-likeness for all stimuli was rated between *unlike a real word* and *neutral*. Across the 12 nonword stimuli, lexical, linguistic, and phonological values remained low, and word-likeness scores for stimuli with initial and noninitial stress did not significantly differ (see Table 2).

Experimental block design. Twelve experimental blocks were created, one for each target nonword. Each block included one nonword target (bisyllabic $C_1VC_2C_3VC_4$) and three monosyllabic CVC nonword foils. Inclusion of three foils allowed target phonemes to appear in at least two nonwords within a block and prevent process of elimination and/or "anticipation" strategies between presentation of the phoneme to be monitored and the nonword cue. Monosyllabic foils were chosen to prevent potential priming of the syllabic stress pattern. Together, the three nonword foils contained all six target phonemes. To avoid any additional potential priming effects, a phoneme that occupied the C_1 or C_3 position in a target nonword did not occupy the onset position of any of the three foils. Likewise, C_2 and C_4 phonemes of the target word did not occupy the final phoneme of any of the three foils, and vowels within foils did not overlap with vowels in target nonwords. See Table 2 for list of nonword–foil blocks.

Stimuli recording. Nonword and phoneme stimuli were recorded by a female native monolingual Standard American English speaker with no history of speech, language, or hearing disorders from eastern Washington state (North American Western dialect: Labov, Ash, & Boberg, 2006). Stimuli were recorded in a sound-treated room with a microphone placed approximately 12 inches from the speaker's mouth using KayPENTAX Computerized Speech Lab (Model 4150; KayPENTAX, Lincoln Park, NJ). Digital files were sampled at a 22050-Hz sampling rate and 16-bit quantization. Stimuli for the tone series were created using an open-source, online sine tone generator (AudioCheck; http://www.audiocheck.net) and compiled using a digital editing program (Audacity, Version 1.2.6; http://www.audacity.sourceforge.net). Phonemes and tones were presented with a mean amplitude of 45 dB SPL.

Procedure and Tasks

Participants completed the experimental portion of the study during the second 90-minute session. The second session included completion of 12 experimental blocks, one for each target nonword. During this session, participants were comfortably positioned in a chair approximately 18 inches from the screen wearing headphones, with fingers resting on preassigned yes and no buttons on the keyboard. Manual responses for the baseline reaction time

Table 2. Lexical, linguistic, and phonological properties of target stimuli with associated foils.

Block	Experiment 1: Initial stress	Experiment 2: Syllable boundary	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/
Factor	**Mean or value**	***Z* score range**			
Word-likeness[a]					
Initial stress (Experiment 1)	2.62	−1.93 to 1.20			
Noninitial stress (Experiment 2)	2.75	−1.53 to 1.53			
Word shape	CVCCVC				
Phonotactic probability—segmental[b]	0.18	−0.20 to 1.24			
Phonotactic probability—biphone[b]	0.01	−1.43 to 1.20			
Neighborhood frequency[c]	0.00				
Neighborhood density[c]	0.00				
Phonetic complexity[d]	6.17	−1.13 to 1.78			
First syllable frequency[e]	1.00	0.00 to 0.00			
Second syllable frequency[e]	0.08	−0.29 to −0.30			
Orthographic transparency[f]	1:1				
Uniqueness point[g]	Third phoneme				
Syllable boundary clarity[h]	10 of 12 clusters illegal in onset and offset				

Note. Values determined using database and/or criteria provided in the following literature:

[a]Word-Likeness Scale (Gathercole, 1995); [b]Phonotactic Probability Calculator (Vitevitch & Luce, 2004); [c]Hoosier Mental Lexicon (Luce & Pisoni, 1998); [d]Word Complexity Measure (Stoel-Gammon, 2010); [e]CELEX database (Baayen, Piepenbrock, & Gulikers, 1995); [f]Damian & Bowers (2003); [g]British English Example Pronunciation Dictionary (accessed via FONRYE English Dictionary, Phonetic and Syllable Search; Mlinar, 2010); [h]Treiman & Zukowski (1990).

task, nonlinguistic auditory monitoring task, and silent phoneme monitoring task were recorded using stimulus presentation software (SuperLab Pro 4.5; Cedrus Corporation, San Pedro, CA). Verbal responses were recorded using a Zoom Q2 audio/video recorder (Zoom Corporation, Tokyo, Japan).

Stimuli Presentation

Each block included a training task followed by the silent phoneme monitoring task. The three-phase training task was completed to teach participants to silently generate target nonwords upon presentation of a neutral nonorthographic visual cue during the experimental task. Training was critical to the present study for at least three reasons. First, as reported by Smits-Bandstra (2010), significant between-groups differences would be expected on the basis of simple motor responses (i.e., button press) if limited practice was provided, regardless of experimental condition. Second, due to the novelty of the stimuli, sufficient practice was necessary to ensure greater accuracy when recalling nonwords (see Gupta, 2003; n = 18 exposures for bisyllabic nonwords; 97.31% accuracy). On the basis of Gupta's (2003) data, within each block, each participant received a minimum of 18 exposures to each target nonword during the training task prior to completing the experimental silent phoneme monitoring task. The third and perhaps most crucial need for this training is that it allowed us to confirm that the participants were able to accurately identify and generate the target nonwords when presented with a nonorthographic cue for these nonwords in the location on the screen they had been trained to associate with that particular nonword. Thus, training was critical because it established the reliable generation of a novel word during silent phoneme monitoring and removed the potential confounds related to "visual scanning" of orthographic representations (e.g., Hawelka, Huber, & Wimmer, 2006; Mason, 1982).

Training task. The training task used was a modified version of the three-phase training procedure described by Levelt and colleagues (e.g., Cholin et al., 2011; Cholin, Levelt, & Schiller, 2006; Levelt & Wheeldon, 1994): (a) repetition phase; (b) identification phase; and (c) generation phase. Description of each phase during the training task sequence is provided in Table 3.

During the repetition phase, participants repeated each of the four auditory nonwords in the block (i.e., one target nonword, three foil nonwords) presented aurally simultaneously with a visual orthographic representation in one of the four corners of the computer screen, with the stressed syllable capitalized (e.g., Experiment 1: "MAZfoov";

Table 3. Structure of a training task and silent phoneme monitoring task within a single block.

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 two exposures to each nonword per phase							
5. If participant failed to accurately produce word two out of two times during final generation phase, repeat Step 4							
6. If participant failed to accurately produce word two out of two 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, or ʃ]	900 ms before visual cue	Neutral icon, Corner A	Corner A	1. Press yes/no 2. Say word	Fixed random	6 (4 yes, 2 no)
Foil 1	[v, m, z, or l]	900 ms before visual cue	Orthographic, Corner B	Corner B	1. Press yes/no 2. Say word	Fixed random	4 (2 yes, 2 no)
Foil 2	[z, f, m, or ʃ]	900 ms before visual cue	Orthographic, Corner C	Corner C	1. Press yes/no 2. Say word	Fixed random	4 (2 yes, 2 no)
Foil 3	[ʃ, l, v, or f]	900 ms before visual cue	Orthographic, Corner D	Corner D	1. Press yes/no 2. Say word	Fixed random	4 (2 yes, 2 no)

Note. Three-phase training procedures similar to paradigms described by Levelt and colleagues (Cholin et al., 2006, 2011; Levelt & Wheeldon, 1994) and Gupta (2003).

Experiment 2: "mazFOOV"). Each nonword within the repetition phase (one target, three foils) was presented four times in randomized order, resulting in 16 total responses (i.e., four exposures to target nonword, four exposures for each of the three foils; see Table 3, Step 1, for detailed description).

During the identification phase, participants heard auditory presentations of the same four stimuli and were instructed to point to the corner of the screen associated with the nonword. During this phase, the orthographic target nonword was replaced with a neutral visual icon (i.e., 2 × 2-inch speaker icon) and presented simultaneously with three orthographic foils in the remaining three corners of the screen. Again, each of the four words within the identification phase was presented aurally four times in random order, resulting in an additional 16 total responses (i.e., four exposures to target nonword, four exposures for each of the three foils; see Table 3, Step 2, for detailed description).

During the generation phase, participants saw either an orthographic foil or the visual icon appear without auditory input. Each appeared in the same designated corners established during the repetition and identification phases. Participants were instructed to say aloud the word associated with the corner of the screen. As in the repetition and identification phase, each of the four words was presented four times in random order, resulting in an additional 16 total responses (i.e., four exposures to target nonword, four exposures for each of the three foils; see Table 3, Step 3, for detailed description).

The entire training task was then repeated prior to the experimental task with fewer exposures. Each of the four nonwords was presented two times, rather than four times, in each of the three phases during this secondary round of training (i.e., six target nonword exposures, six exposures for each of the three foils; see Table 3, Step 4). Together, the initial and secondary training resulted in 72 total exposures (i.e., 18 for the target nonword, 18 for each of the three foils). Portions of training that required verbal responses were self-paced to accommodate for possible disfluent responses. If the participant could not identify the target nonword with 100% accuracy during the generation phase of the second round of training, this second round of training sequence (Table 3, Step 4) was repeated until 100% response accuracy was achieved (Table 3, Steps 5 and 6).

Silent phoneme monitoring task. The experimental silent phoneme monitoring task was completed after the training task for each block, as depicted in Table 3. Participants were instructed: "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 then be cued to say the word aloud. Press Enter to get started." Each silent phoneme monitoring trial occurred in the sequence depicted in Figure 1.

The silent phoneme monitoring task within each block consisted of 12 trials; each phoneme was heard twice. Phonemes to be monitored occurred in one of four consonant positions within target nonword stimuli (i.e., C_1, C_2, C_3, or C_4) or one of three foils. All phonemes received both a yes and no response to prevent anticipation or guessing from the participant. Target nonwords appeared six times (four yes responses for each C_1 to C_4 position, two no responses for phonemes not present). The sequence of trials within blocks were presented in a fixed randomized structure to ensure that (a) the same target phonemes were not presented on consecutive trials, (b) the same nonword was not presented on consecutive trials, and (c) no distinguishable yes/no pattern could be detected. Once all 12 trials within a block were completed, the next experimental block began.

To reduce visual habituation or expectancy, each target nonword (represented by the visual icon) appeared in each of the four corners of the screen three times across all 12 blocks. For each consonant position (C_1, C_2, C_3, C_4), all 22 participants had the opportunity to provide 12 yes tokens (i.e., true positive), for a total of 132 true positive tokens per consonant position considered during final analyses. Presentation order of the 12 experimental blocks was randomized between participants, and a 15-minute break was provided after the sixth block to avoid fatigue.

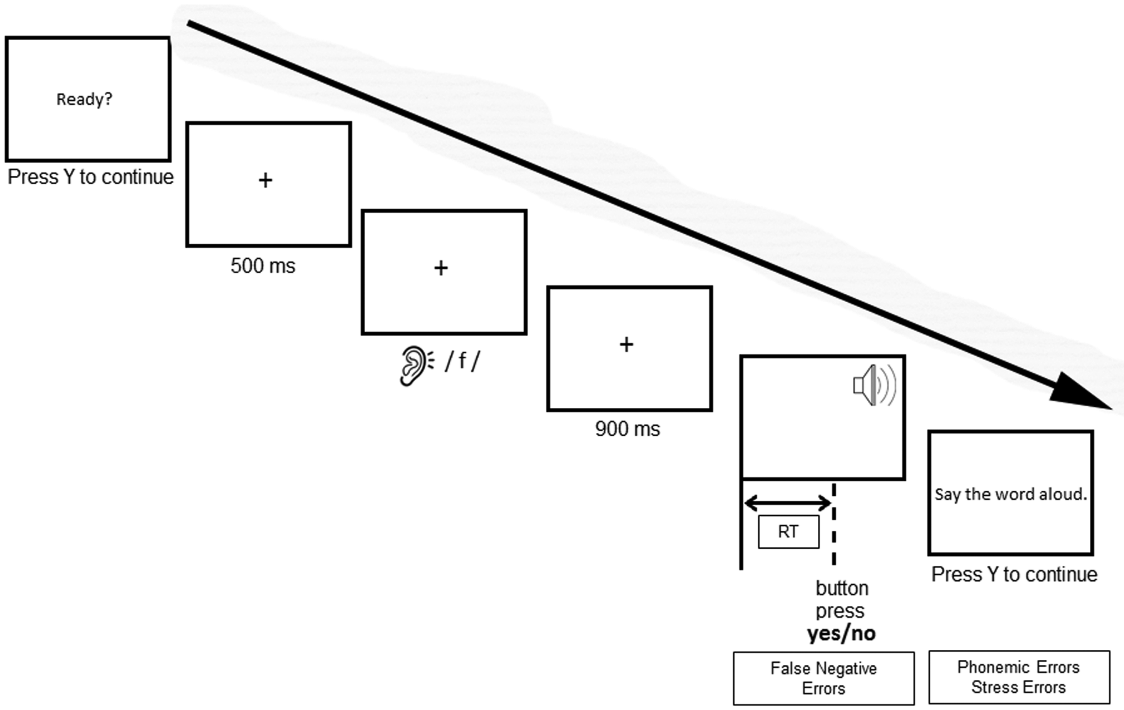
Manual reaction time responses were recorded as the duration between the onset of the speaker icon until the yes or no button was pressed (see Figure 1). After either button was pressed or 3,000 ms had elapsed, the participant was prompted to say the word aloud and press yes to continue, which triggered the beginning of the next trial starting at the "Ready?" screen. This final screen was self-paced, and accuracy of participants' verbal production was recorded by the examiner and video-recorded to ensure the appropriate nonword had been generated.

Analyses

The purpose of Experiment 1 was to determine the influence of the metrical property of initial stress on the speed and accuracy of phonological encoding in AWS and AWNS. To assess reaction time latencies for nonwords with initial stress, a 2 × 4 repeated measures mixed-model analysis of variance (ANOVA) was conducted, with Talker Group (i.e., AWS, AWNS) as the between-groups factor, and Phoneme Position Latencies (i.e., C_1–C_2–C_3–C_4) as the within-group factor. Planned pairwise comparisons were conducted to examine mean latencies of consonants flanking initial stress (i.e., C_1, C_2) and syllable boundary (i.e., C_2, C_3) within and between talker groups. False negative errors were also examined using a 2 × 4 repeated measures mixed-model ANOVA, with Talker Group as the between-groups factor and Phoneme Position Error Rate as the within-group factor. To examine posttrial verbal error data between groups, two separate one-way analysis of covariances (ANCOVAs) were conducted for each error type (i.e., phonemic errors and stress errors), with talker group as the independent variable, error rate as the dependent variable, and Participant as the covariate factor to control for individual differences.

Individual alpha levels were set at .05 for all ANOVA, ANCOVA, and pairwise comparisons. Independent *t* tests

Figure 1. Reaction time (RT) latency and accuracy measurements during adapted silent phoneme monitoring task. Errors include false positive responses, phonemic errors during posttrial verbal response, and stress-based errors during posttrial verbal response.



were conducted to compare group performance on age, baseline reaction time, and phonological processing subtest performance. In addition, independent Pearson correlation analyses were conducted between experimental data (i.e., monitoring latencies, false negative errors, phonemic errors, and stress errors) and performance on phonological processing subtests, as well as age, stuttering severity score, baseline reaction time, and nonlinguistic auditory monitoring for each talker group. 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.

Results

Age, baseline reaction time, and phonological subtest scores. As depicted in Table 4, independent t tests found no significant differences in age or baseline reaction time. In addition, no group differences were observed in eight of the nine phonological subtests. However, there was a significant talker group difference in onset generation of real words, $t(20) = 2.34$, $p = .024$, with AWNS exhibiting slightly better segmentation skills ($M = 19.73$, $SD = 0.47$) than AWS ($M = 18.82$, $SD = 0.87$). Both groups performed at or near ceiling and mean variability was small; thus, this difference was not considered to be meaningful in nature (for a similar argument, see Sasisekaran et al., 2006, p. 14).

Forward digit span was negatively correlated with false negative error rate ($r = -.616$; $p = .044$) and phonemic error rate ($r = -.620$; $p = .042$) for AWS, but not AWNS (false negative: $r = -.269$; $p = .424$; phonemic: $r = .161$; $p = .636$). All remaining factors were not significantly correlated with experimental data, including onset generation scores (AWNS: $r = .124$; $p = .716$; AWS: $r = .053$; $p = .876$; see Figure 2). Based on these findings, forward digit span performance was included as a covariate during analyses of false negative error and phonemic error rate, with reported means adjusted for the effect of the covariate.

Nonlinguistic auditory monitoring task. Participant responses during the nonlinguistic auditory monitoring task were assessed using a mixed-model repeated measures ANOVA, with Talker Group as the between-groups factor (i.e., AWNS, AWS) and Tone Position (i.e., positions 1–4) as the within-group factor. False negative responses were defined as a manual response of no when the high tone was present (i.e., true positive). Reaction times that exceeded two standard deviations above or below the participant's mean during accurate responses for each position were considered outliers. Of the 1,056 true positive responses, 5.68% ($n = 60$) were false negative responses (AWNS, $n = 33$; AWS, $n = 27$), and 5.02% ($n = 53$) were outliers (AWNS, $n = 25$; AWS, $n = 28$). False positive responses and outliers were excluded from analyses. Final reaction time analysis was conducted on 943 viable responses (i.e., AWNS, $n = 470$; AWS, $n = 473$).

A significant main effect for Reaction Time × Position was found, $F(3, 18) = 62.00$, $p < .001$, $\eta\rho^2 = .756$, but no main effect for group ($F < 1$) or interaction ($F < 1$). Post

Table 4. Summary of participant demographics and screening measures.

Characteristic	Experiment 1: Initial stress			Experiment 2: Syllable boundary		
	AWNS	AWS	*p*	AWNS	AWS	*p*
n (male, female)	11 (7, 4)	11 (7, 4)		11 (7, 4)	11 (7, 4)	
Age (y)	20.36 (2.62)	22.27 (6.45)	.374	23.45 (4.97)	22.27 (2.97)	.506
Baseline RT (ms)	318.65 (32.63)	323.14 (44.19)	.789	319.30 (25.67)	336.43 (38.74)	.238
Nonword Repetition[a]	12.80 (2.04)	12.36 (1.96)	.624	12.18 (1.47)	12.27 (1.74)	.896
Onset ID: *Real Words*[a]	13.82 (2.56)	13.00 (3.61)	.546	11.91 (3.39)	12.36 (3.61)	.764
Onset ID: *Nonwords*[a]	15.91 (3.39)	14.18 (4.09)	.294	12.82 (4.22)	12.00 (4.75)	.674
Onset Generation: *Real Words*[b]	19.73 (.47)	18.82 (.87)	.006*	19.36 (.92)	19.36 (.81)	1.000
Onset Generation: *Nonwords*[b]	19.55 (.69)	19.45 (1.21)	.831	19.36 (.67)	19.55 (.52)	.488
Rhyme ID[c]	14.91 (.30)	14.82 (.60)	.660	14.91 (.30)	14.91 (.30)	1.000
Rhyme Generation[c]	14.55 (.69)	14.36 (.67)	.538	14.27 (.91)	14.64 (.67)	.298
Forward Digit Span[d]	11.00 (1.90)	10.82 (2.27)	.841	8.82 (1.25)	11.09 (2.26)	.010*
Backward Digit Span[d]	8.18 (2.32)	6.55 (2.162)	.102	5.82 (1.72)	6.82 (1.47)	.159

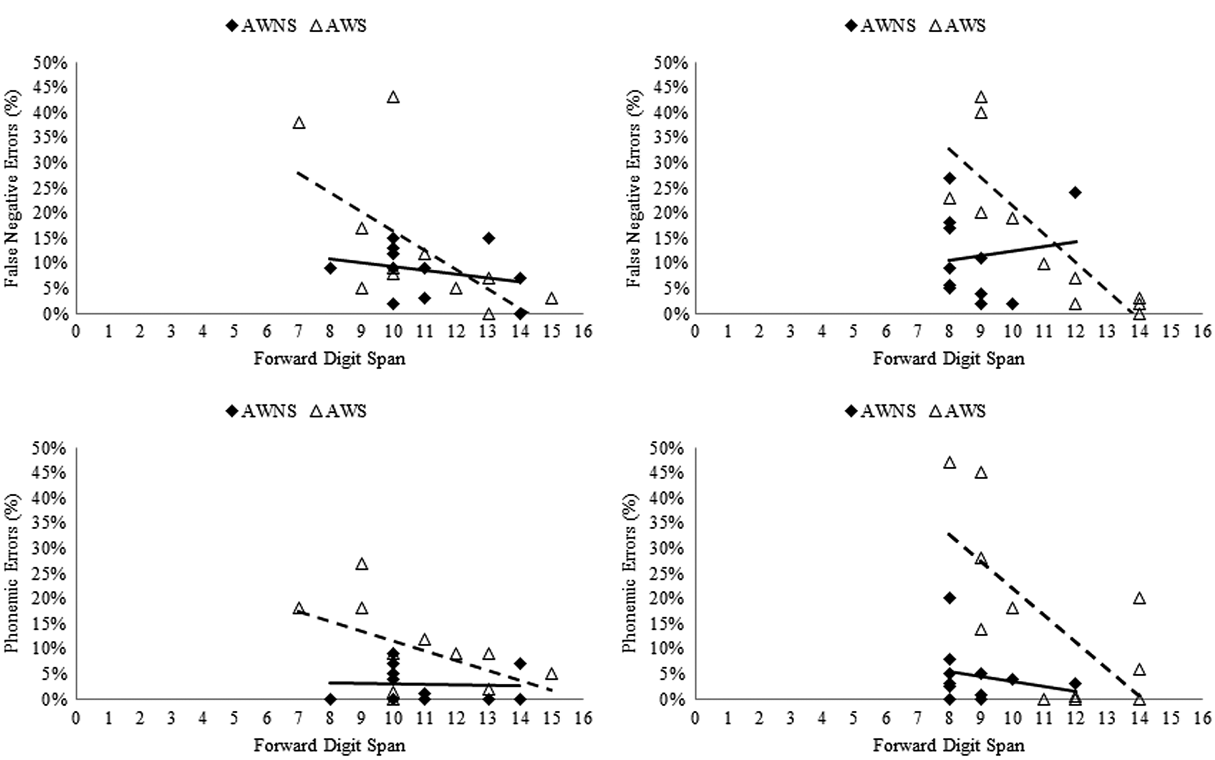
Note. Except for *n* (male, female), data are presented as mean (standard deviation); AWNS = adults who do not stutter; AWS = adults who stutter; RT = reaction time; ID = identification.

[a]Comprehensive Test of Phonological Processes (CTOPP) subtest (Wagner et al., 1999; ceiling–20); [b]CTOPP subtest–adapted (Wagner et al., 1999; ceiling–20); [c]Sasisekaran & Byrd (2013; ceiling–15); [d]Wechsler Adult Intelligence Scale–Third Edition (Wechsler, 1997; Forward Digit Span ceiling–16; Backward Digit Span ceiling–14).

*$p < 0.05$.

hoc tests revealed that tones were identified at significantly longer latencies for both groups as series position increased, although latencies between Position 1 and Position 2 did not reach significance. Error analysis was conducted using the 60 false negative errors (AWNS: $n = 33$; AWS: $n = 27$). Similar to reaction time data, a significant main effect emerged for position, $F(3, 18) = 9.75$, $p < .001$, $\eta\rho^2 = .328$, but there was no main group effect ($F < 1$) or interaction ($F < 1$). Post hoc comparisons revealed nonsignificant differences between groups, with only AWNS identifying tones at Position 3

Figure 2. Scatter plots of forward digit span performance and mean percentage of false negative errors (top) and phonemic errors (bottom) for adults who do and do not stutter (AWNS and AWS, respectively) during a silent phoneme monitoring task of $C_1VC_2C_3VC_4$ nonwords with initial stress (Experiment 1, left) or initial syllable boundary (Experiment 2, right).



and Position 4 with significantly lower accuracy than Position 1, which was identified with 100% accuracy.

Silent phoneme monitoring. Similar to Sasisekaran et al. (2006), data were removed from reaction time and error analyses and considered unusable if the following criteria were met:

- No response: participant provided no manual response or initiated fluent verbal response more than 3,000 ms after button press
- Overlapping verbal response: verbal response overlapped manual response
- Outlier: manual response exceeded two standard deviations above or below the participant's consonant position–specific mean reaction time for fluent, accurate responses

In sum, from the initial 1,056 tokens collected, 110 tokens (10.42%) were considered unusable prior to all subsequent analyses. The final data corpus included 946 usable tokens (AWNS, $n = 453$; AWS, $n = 493$).

To reduce the likelihood that reaction times were affected by inaccurate retrieval, processing (i.e., identification of phonemes within the target nonword) of reaction time data was on the basis of accurate nonverbal monitoring accompanied by accurate, fluent verbal responses. Individual tokens were excluded from the reaction time analysis if they met the following criteria:

- False negative: participant manually responded no when target phoneme was present
- Phonemic error: posttrial verbal response included one or more phonemic error
- Stress error: posttrial verbal response included inaccurate syllabic stress
- Stuttered response: posttrial verbal response contained a stuttering-like disfluency (i.e., sound-syllable repetition, audible and inaudible sound prolongations)

From the usable 946 tokens collected (AWNS, $n = 453$, AWS, $n = 493$), 53 tokens were excluded from AWNS participants (false negative, $n = 39$ [8.60%]; phonemic error, $n = 11$ [2.43%]; stress error, $n = 3$ [0.66%]; stuttered response, $n = 0$ [0.00%]) and 121 from AWS participants (false negative, $n = 66$ [13.38%]; phonemic error, $n = 31$ [6.28%]; stress error, $n = 2$ [0.41%]; stuttered response, $n = 22$ [4.46%]). In total, 174 tokens (18.39%) were excluded from the reaction time analysis on the basis of error response and/or disfluent posttrial verbal response. The final data corpus included 772 fluent, accurate tokens (AWNS, $n = 400$; AWS, $n = 372$).

Latencies. A mixed-model repeated measure ANOVA was conducted to assess the reaction time latencies of AWNS and AWS during silent phoneme monitoring of consonants flanking initial stress assignment in bisyllabic nonwords (i.e., C_1 and C_2). There was no significant Group × Position interaction ($F < 1$), and no significant main effect for talker group ($F < 1$). However, there was a significant main effect for consonant position, $F(3, 18) = 13.48$, $p < .001$, $\eta\rho^2 = .692$. As depicted in Figure 3, planned within-group comparisons revealed significantly longer latencies for C_2 ($M = 1100.21$, $SE = 102.59$) relative to C_1 ($M = 891.23$, $SE = 81.09$) in AWS as well as AWNS (C_1: $M = 988.00$, $SE = 81.09$; C_2: $M = 1095.64$, $SE = 102.59$), but no significant latency differences were observed between C_2 and C_3. Planned comparisons did not indicate significant mean latency differences between talker groups for any phoneme position.

False negatives. As noted, forward digit span performance was included as a covariate during repeated-measures ANOVA of monitoring latencies. As seen in Figure 3, no significant main effect for position was detected, $F(3, 18) = 1.78$, $p = .182$, $\eta\rho^2 = .085$. No significant between-groups differences at any position were detected, $F(1, 19) = 1.12$, $p = .303$, $\eta\rho^2 = .056$, and there was no significant Group × Position interaction ($F < 1$).

Phonemic and stress errors. No significant main effect of group was revealed for phonemic error, $F(3, 22) = 1.06$, $p = .318$, $\eta\rho^2 = .164$, upon inclusion of forward digit span as a covariate and adjusting for individual participant. Stress assignment errors did not significantly differ between groups ($F < 1$), again after adjusting for individual participant differences. See Figure 4 for an illustration of posttrial verbal errors produced by AWS and AWNS during Experiment 1.

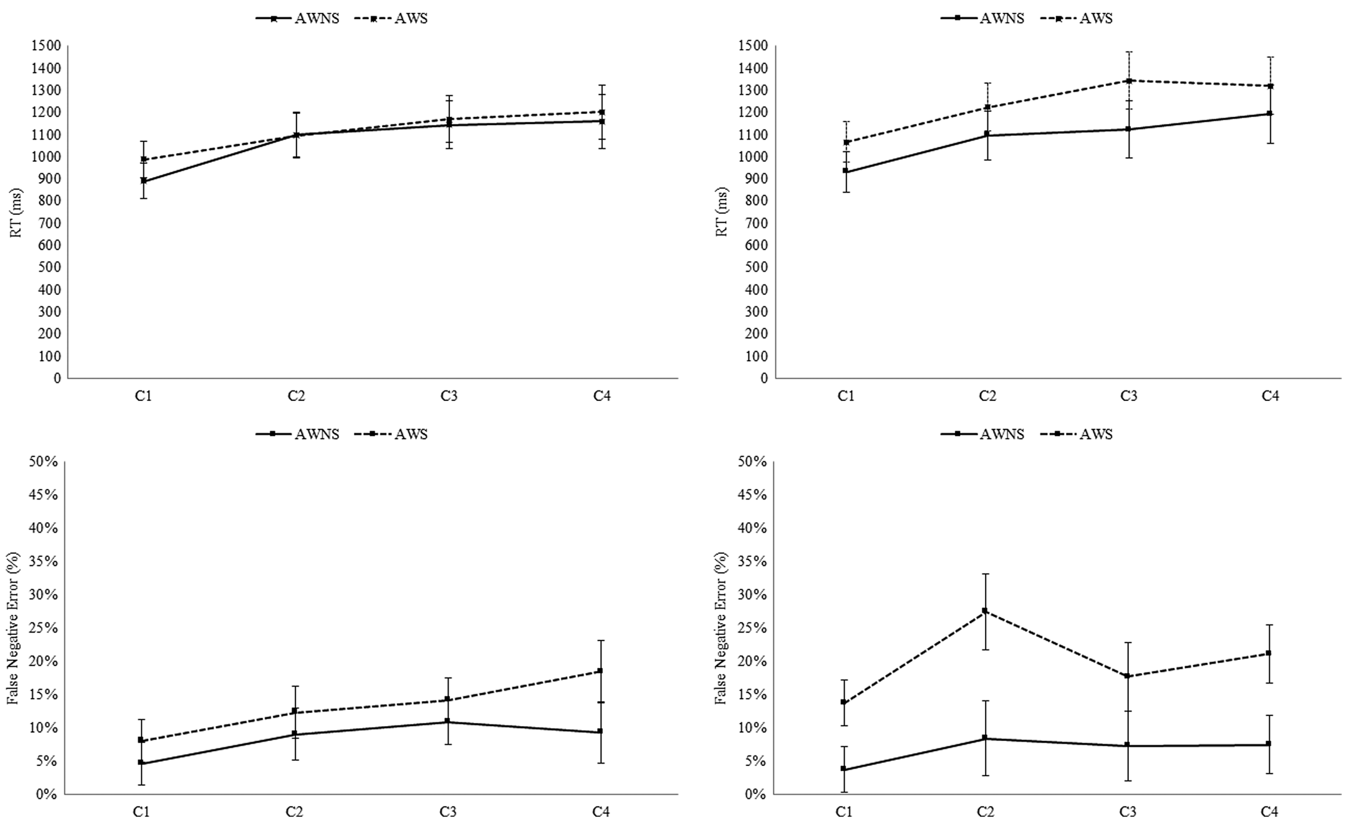
Experiment 2: Influence of Syllable Boundary

Participants

Intake procedure and inclusionary criteria were identical to those of Experiment 1. Experiment 2 also included 22 adults (AWS = 11, seven men, four women, age range: 19 to 28 years, $M = 22$ years; AWNS = 11, seven men, four women, age range: 18 to 36 years, $M = 23$ years). These participants did not participate in Experiment 1. Participants provided informed consent and were compensated for participation.

Stuttering severity. Participant stuttering severity measures were obtained using the same 9-point stuttering severity scale (O'Brian et al., 2004) used in Experiment 1. All 11 AWS participants self-identified as AWS, and eight of 11 had received a prior diagnosis of stuttering. The remaining three of 11 were diagnosed with stuttering by the first author, an ASHA-certified, licensed speech-language pathologist. Each AWS scored 2 or higher on either the conversational or the reading sample. The 11 AWNS participants did not self-identify as an AWS, had no prior diagnoses of stuttering, and received a score of no greater than 1 on conversational and reading samples (see Table 1). Interrater reliability of stuttering severity was determined by the first author and the same undergraduate research assistant who conducted Experiment 1. For AWS, interrater reliability for both speech samples was within 1 point for 100% of the randomly selected participants ($n = 3$), with all participants receiving a score of 2 or higher (ICC = .93 for conversation, .80 for reading). Intrarater reliability was within 1 point for 100% of the participants (ICC = 1.00 for

Figure 3. Mean reaction time latencies (RT, top row) and mean percentage of false negative errors (bottom row) for adults who do not stutter (AWNS) and adults who stutter (AWS) during a silent phoneme monitoring task of $C_1VC_2C_3VC_4$ nonwords with initial stress (Experiment 1, left) or initial syllable boundary (Experiment 2, right). Error bars represent standard error of the mean. False negative errors reflect adjusted means upon inclusion of forward digit span as covariate.



conversation, .93 for reading). For conversational and reading samples, severity ratings for AWNS participants was the same for all three randomly selected participants; each participant received a score of 1 (i.e., *no stuttering*).

Experimental Design, Procedure, and Analyses

Experiment 2 was identical to Experiment 1 in experimental design, procedures, stimuli recording protocol, and phonetic structure of stimuli. The critical difference between the two experiments was the assignment of stress within the nonword stimuli. As in Schiller (2005), stress was assigned to the second syllable of the $C_1VC_2C_3VC_4$ bisyllabic nonword. Thus, syllable boundary assignment was the first-encountered metrical property during phonological encoding within the silent phoneme monitoring task.

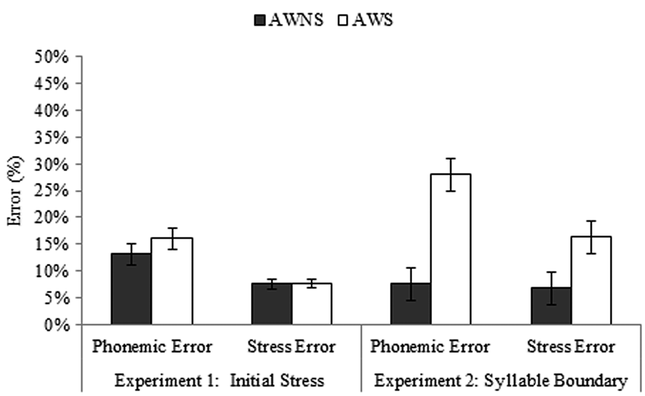
Results

Age, baseline reaction time, and phonological subtest scores. As depicted in Table 4, independent *t* tests found no significant differences in age or baseline reaction time. In addition, no group differences were observed in eight of the nine phonological subtests. However, talker groups demonstrated a significant difference in forward digit span, $t(20) = -2.92$, $p = .010$), with AWS exhibiting slightly higher mean performance ($M = 11.09$, $SD = 2.26$) than AWNS ($M = 8.82$, $SD = 1.25$).

Similar to Experiment 1, forward digit span was negatively correlated with false negative error rate ($r = -.829$; $p = .002$) and phonemic error rate ($r = -.728$; $p = .011$) for AWS but not AWNS (false negative: $r = .142$; $p = .678$; phonemic: $r = -.306$; $p = .360$). Figure 2 indicates greater variance in AWS performance in Experiment 2 (range = 8 to 14, with three of 11 scoring 14) than AWNS (range = 8 to 12, with 10 of 11 scoring below 10). All remaining factors were not significantly correlated with experimental data. Based on these findings, forward digit span performance was included as a covariate during analyses of false negative error rate and phonemic error rate with reported means adjusted for the effect of the covariate.

Nonlinguistic auditory monitoring task. Participant responses during the nonlinguistic auditory monitoring task were assessed using a mixed-model repeated measures ANOVA, with Talker Group as the between-groups factor and Tone Position as the within-group factor. Of the 1,056 true positive responses, 4.73% ($n = 50$) were outliers (i.e., AWNS, $n = 23$; 2.56% AWS, $n = 27$), and 3.41% ($n = 36$) were false negative responses (i.e., AWNS, $n = 24$; AWS, $n = 12$). False negative and outlier responses were excluded

Figure 4. Mean percentage of posttrial phonemic errors and stress errors by adults who do not stutter (AWNS) and adults who stutter (AWS) during a silent phoneme monitoring task of $C_1VC_2C_3VC_4$ nonwords with initial stress (Experiment 1, left) or initial syllable boundary (Experiment 2, right). Error bars represent standard error of the mean. Phonemic errors reflect adjusted means upon inclusion of forward digit span as a covariate.



from analyses. Final reaction time analysis was conducted on the remaining 970 viable responses (i.e., AWNS, n = 481; AWS, n = 489). A significant main effect for position was found, $F(3, 18) = 245.00$, $p < .001$, $\eta\rho^2 = .925$, but no main group effect, $F(1, 20) = 2.487$, $p = .130$, $\eta\rho^2 = .111$, or interaction was revealed ($F < 1$). Post hoc comparisons revealed that each tone was identified at significantly longer latencies for both groups as serial position increased. Error analysis was conducted on the 36 viable responses that included the false negative errors excluded from reaction time analysis. Similar to reaction time data, a significant main effect for position was found, $F(3, 18) = 9.98$, $p < .001$, $\eta\rho^2 = .333$, but no main group effect, $F(1, 20) = 2.06$, $p = .166$, $\eta\rho^2 = .094$, or interaction was revealed ($F < 1$). Post hoc comparisons revealed nonsignificant between-groups differences in accuracy, with AWNS identifying tones at Position 4 (M = 11.82%, SE = 2.73%) with significantly lower accuracy than Position 2 (M = 0.82%, SE = 0.00%) and Position 1, which identified tones with 100% accuracy.

Silent phoneme monitoring. The purpose of Experiment 2 was to assess reaction time latencies and accuracy of AWS and AWNS during silent monitoring of phonemes at the syllable boundary of nonwords. Similar to Experiment 1, data were removed from reaction time and error analyses and considered unusable if the participant provided no response, overlapping verbal response, or manual response considered an outlier. From the initial 1,056 tokens collected (AWNS, n = 528; AWS, n = 528), 68 tokens (6.44%) were considered unusable prior to analyses. The final data corpus included 988 usable tokens (AWNS: n = 500; AWS: n = 488).

Latencies. The following errors were removed from reaction time analysis: false negatives, phonemic errors, stress errors, and stuttered responses. From the usable data corpus, 80 tokens were excluded from AWNS participants (false negative, n = 59 [11.80%]; phonemic error, n = 16 [3.20%]; stress error, n = 5 [1.00%]; stuttered response, n = 0 [0.00%]) and 156 from AWS participants (false negative, n = 78 [15.98%]; phonemic error, n = 42 [8.61%]; stress error, n = 28 [5.74%]; stuttered response, n = 8 [1.64%]). In total, 236 tokens (23.89%) were excluded on the basis of error response and/or disfluent posttrial verbal response, resulting in a final data corpus that included 752 fluent, accurate tokens (AWNS: n = 420; AWS: n = 332).

A mixed-model repeated measures ANOVA was conducted to assess the reaction time latencies of AWNS and AWS during silent phoneme monitoring of phonemes flanking the syllable boundary (i.e., C_2 and C_3). No significant Group × Position interaction was revealed ($F < 1$) nor was a significant main effect for talker group found ($F < 1$). However, there was a significant main effect for consonant position, $F(3, 18) = 9.82$, $p < .001$, $\eta\rho^2 = .621$. Planned comparisons indicated mean latencies between talker groups did not significantly differ at any consonant position (i.e., C_1 to C_4). However, planned within-group comparisons revealed significantly longer latencies between C_2 (M = 1223.72, SE = 108.35) and C_3 (M = 1342.34, SE = 129.00) in the AWS group. AWNS did not exhibit a significant difference between C_2 and C_3 (C_2: M = 1095.71, SE = 108.35; C_3: M = 1122.50, SE = 129.00). Significant differences were also observed between C_1 and C_2 latencies for both AWS (C_1: M = 931.33, SE = 89.91; C_2: M = 1095.71, SE = 108.35) and AWNS (C_1: M = 1066.88, SE = 89.91; C_2: M = 1223.72, SE = 108.35).

False negatives. Forward digit span performance was included as a covariate during analysis. No significant differences by position ($F < 1$) or Group × Position interaction ($F < 1$) were indicated. However, a significant between-groups difference was detected, $F(1, 19) = 6.30$, $p = .021$, $\eta\rho^2 = .249$, with AWS exhibiting significantly greater false negative errors. Post hoc comparisons revealed an overall greater error rate across positions for AWS relative to AWNS, but group differences reached significance for the C_2 position (AWS: M = 27.45%, SE = 5.65%; AWNS: M = 8.39%, SE = 5.65%). Figure 3 depicts adjusted mean values of false negative errors after inclusion of covariate.

Phonemic and stress assignment errors. Upon inclusion of forward digit span and participant as covariates, a significant main effect between groups was found, $F(1, 21) = 18.85$, $p < .001$, $\eta\rho^2 = .512$, with AWS exhibiting greater phonemic errors (M = 27.95%, SE = 3.06%) than AWNS (M = 7.56%, SE = 3.06%). A significant main effect was also found between groups, $F(1, 21) = 5.17$, $p = .035$, $\eta\rho^2 = .214$, with AWS exhibiting greater stress assignment errors (M = 16.37%, SE = 2.97%) than AWNS (M = 6.78%, SE = 2.97%; see Figure 4) during the posttrial verbal response.

Discussion

The primary goal of this study was to examine the extent to which metrical encoding differs in AWS. Similar to Schiller's (2005) research, stress and syllable boundary were isolated in two separate experiments to allow for independent examination of each as the first-encountered metrical property. Three main findings were observed. First, AWS and AWNS demonstrated comparable speed and accuracy when monitoring phonemes in stimuli with initial stress. Second, in the absence of initial stress, only AWS exhibited significantly delayed monitoring of phonemes after the syllable boundary. Third, when silently monitoring stimuli with noninitial stress, AWS exhibited significantly more errors than AWNS when identifying phonemes immediately before the syllable boundary as well as significantly more phonemic and stress-assignment errors during posttrial production. Together, findings suggest that initial stress assignment does not alter the phonological encoding patterns in AWS in a manner distinct from AWNS. However, in the absence of initial stress, speed and accuracy of phonological encoding at the syllable boundary appears to be uniquely compromised for AWS.

Initial Stress

When monitoring phonemes in syllables with initial stress both talker groups demonstrated significant C_1–C_2 latencies that have been documented in past research (e.g., Sasisekaran et al., 2006; Schiller, 2005). The presence of initial stress did not result in a significantly greater delay between C_1 and C_2 encoding for AWS as compared with AWNS. Findings suggest that the time course of phonological encoding in AWS is not disproportionately slowed when assigning metrical stress to the initial syllable.

Data from Experiment 2 warrant closer examination of the role of initial stress during segmental encoding. Although left-to-right incremental encoding was maintained across experiments, significant C_1–C_2 latencies were also observed for both groups in the absence of initial stress. In contrast to previous findings (Wheeldon & Levelt, 1995; Wheeldon & Morgan, 2002; Schiller, 2005), our data suggest delayed segmental encoding within the initial syllable may not be exclusively linked to stress assignment. However, the unique methodology of the present study may account for these unexpected findings.

To our knowledge, the present study was the first to use nonwords in a silent phoneme monitoring task. According to Cholin et al. (2006), phonological code comprising low-frequency and/or nonword syllables do not have access to the "mental syllabary" described by Levelt et al. (1999) as a fast-access repository for high-frequency syllables. Instead, when encoding nonwords, segmental information must be fully computed from beginning to end in a more laborious, incremental fashion before it becomes available to the articulatory system. Perhaps the use of nonwords in the present study restricted encoding to this more "indirect" route across experiments regardless of stress assignment.

That being said, if present, task complexity differences that are based on nonlexical stimuli did not affect the talker groups equally. The AWNS who participated in the present study appear to monitor phonemes in nonwords (Experiment 1, C_1 to C_4: 891.23 to 1,159.91 ms) more slowly than real-word stimuli in previous studies (Sasisekaran et al., 2006; C_1 to C_4: approximately 775 to 1,050 ms, on the basis of visual inspection of Figure 2, p. 12). In contrast, AWS monitored phonemes within nonwords (Experiment 1, C_1 to C_4: 988 to 1,201 ms) at speeds comparable to real words in previous studies (e.g., Sasisekaran et al., 2006: approximately 975 to 1,300 ms). These findings introduce the intriguing possibility that AWS may use a similar process in their planning for well-learned and novel words. Cholin et al. (2006) state that this indirect route is typically used to prepare speech in situations that require more online, conscious control of verbal production (e.g., lectures, self-corrections). Considering that AWS have a lifelong history of stuttering, a more conscious or less automatic planning for speech production may, in fact, be the norm. Future investigations are warranted to compare speed of segmental encoding for lexical and nonlexical stimuli between groups.

Similar rates of identification of target phonemes in the presence or absence of initial stress (Experiments 1 and 2) are consistent with data reported in Burger and Wijnen's (1999) implicit priming study using real-word stimuli. In their study, AWS and AWNS repeated word sets that shared initial CV segments with and without initial stress. They predicted AWS would exhibit faster speech onset latencies than AWNS for word lists that shared initial stressed CV segments due to inherent difficulty processing the initial stress-bearing vowel but perform similarly to AWNS for unstressed CV stimuli. Instead, both talker groups benefited equally from repetition of CV-matched words, irrespective of stress assignment. These data, along with data from Experiments 1 and 2, suggest that observed C_1–C_2 differences found in both AWS and AWNS may not be attributed solely to online processing of initial stress.

As has been suggested by Smith and colleagues (e.g., Smith, Sadagopan, Walsh, & Weber-Fox, 2010), there may be a critical interplay between phonological encoding and motor programming in AWS. Natke et al. (2002) found that although increased stuttering occurred with greater frequency on stressed syllables relative to unstressed syllables in the onset position, significantly greater stuttering was found for stressed syllables with short vowel duration compared with unstressed onset syllables. In contrast, initial stressed syllables with longer vowel duration did not significantly differ from initial, unstressed syllables. Motoric programming of initial stress composed of short phonetic duration may reduce the time available to completely encode the upcoming syllable and, perhaps, more critically, the upcoming syllable boundary. Thus, the assignment of the initial syllable boundary, rather than the assignment of initial stress, may be a more pivotal factor when assessing the relationship between metrical properties and stuttered speech.

Initial Syllable Boundary

In contrast to Experiment 1, AWS demonstrated a distinctly different within-group phoneme monitoring pattern when the syllable boundary was the first-encountered metrical property. AWS were significantly slower identifying phonemes after syllable boundary assignment (C_2–C_3 latency difference: 118.62 ms) than AWNS (C_2–C_3 latency difference: 24.79 ms). At minimum, these data suggest that assignment of the syllable boundary in the absence of initial stress may uniquely delay phonological encoding in AWS. However, contrary to our expectations, these difficulties emerged even in the presence of significant C_1–C_2 latency differences. Data from Experiment 1 and 2 suggest that the syllable boundary assignment may be particularly challenging for AWS when accessing less common metrical patterns.

According to Levelt et al.'s (1999) model of speech production, stored word forms, or *lexemes*, contain both segmental information (i.e., phonemes) and metrical information (i.e., stress assignment and number of syllable boundaries). Levelt et al. further propose that words with high-frequency stress patterns, such as initial stress in English, are not stored as part of the lexeme but encoded as the default metrical pattern during syllabification (p. 22). Only nondefault stress patterns (i.e., noninitial stress) are stored as part of the target word-form representation. If this is the case, syllabification in AWS may be particularly compromised when encoding word forms that include segmental properties (i.e., phonemes) and each metrical property (i.e., noninitial stress and number of syllable boundaries), as observed during Experiment 2. In contrast, processing remained relatively intact when only segmental information and one metrical property (i.e., number of syllable boundaries) were required during encoding, as observed in Experiment 1. Thus, increased difficulty assigning phonemes at the syllable boundary may be evident for AWS only when greater demands are placed on the phonological encoding system.

This interpretation of findings also corresponds with neurophysiological tone-monitoring data in Mandarin-speaking AWS and AWNS. As reported by Chen, Chen, and Dell (2002), *tone assignment* is defined as a metrical property that conveys distinct lexical meaning within the Mandarin language and, thus, is unlikely to rely on default metrical patterns. That is, phonological representations (e.g., /ma/) must be accompanied by metrical information (e.g., rising pitch = *mother*; falling pitch = *scold*) to convey the appropriate lexical concept. In addition, unlike English, Mandarin speakers have to encode both segmental and unique metrical information without the benefit of high-frequency, default stress patterns. Q. Zhang and Damian (2009) found that nonstuttering Mandarin-speaking adults activate this metrical information (i.e., tone assignment) more slowly than segmental information, although both "arrive" during phonological encoding simultaneously. Data reported by J.-J. Zhang and Xiao (2008) suggest that this slowed activation of nondefault metrical information may be further delayed in AWS. In their behavioral tone-monitoring study, Mandarin-speaking AWS were significantly slower than fluent peers when silently monitoring tone-bearing vowels within target words. These findings support patterns we observed in Experiment 2, demonstrating that significantly delayed retrieval of nondefault metrical frames may uniquely compromise the speed and accuracy of rapid syllabification prior to speech production in AWS.[1]

Yet another consideration is that vulnerability at the syllable boundary observed in Experiment 2 may reflect the combined impact of difficulties processing less common segmental and metrical information in AWS. As noted, both segmental and metrical information are required to formulate the phonological code. The majority of the stimuli in the present study (i.e., 10 of 12 nonwords) included medial clusters constructed to maintain syllable boundary clarity—each cluster was illegal in both onset and coda position—and, therefore, lower in biphone frequency (M = 0.0002, SD = 0.0002). Perhaps, due to underspecification of phonological representations in individuals who stutter (e.g., Anderson, 2007; Anderson & Byrd, 2008; Byrd, Vallely, Anderson, & Sussman, 2012), accessing less common metrical frames in combination with low frequency medial clusters in Experiment 2 further delayed encoding at the syllable boundary. Secondary analyses by Wheeldon and Morgan (2002, pp. 516–517) support the potential influence of segmental properties during assignment of the metrical syllable boundary. That is, removal of stimuli with the "clearest" medial clusters, described by authors as clusters illegal in the onset and offset position (i.e., low frequency), reduced the magnitude of previously significant C_2–C_3 latency differences. Thus, the aggregated impact of delayed encoding of both segmental and metrical information in AWS may further impede efficient planning prior to production.

Errors

False negative errors, phonemic errors, and stress-based errors were similar for both groups in Experiment 1. However, AWS in Experiment 2 exhibited significantly poorer accuracy than fluent peers during silent identification of C_2 phonemes in the absence of initial stress. This difference may reflect the nonlexical status of the stimuli during retention of novel word forms. Treiman and Danis (1988) found adults may establish more robust representations for initial C_1 phonemes of nonwords than the subsequent vowel + C_2 coda. Furthermore, Treiman, Fowler, Gross, Berch, and Weatherston (1995) found adults less able to retain C_2 than C_3 consonants when repeating bisyllabic nonwords with noninitial stress, whereas minimal differences were observed for nonwords with initial stress. The presence of these errors in Experiment 2 but not Experiment 1 suggests that default stress patterns may diminish segmental errors in AWS, and less common metrical configurations may increase segmental errors.

[1]Although the speech production model proposed by Levelt et al. (1999) describes metrical encoding in terms of syllabic stress assignment observed in Germanic languages rather than tonal languages, recent modifications to this model proposed by Roelofs (2014) indicate that although word-form differences exist among languages, the underlying principles of the syllabification process are similar across languages.

Given the overall weaker covert identification of individual phonemes for AWS in Experiment 2, it is not surprising that AWS also exhibited significantly greater phonemic errors during posttrial production. However, an unexpected outcome of Experiment 2 was the significant tendency for AWS to assign stress to the first syllable instead of the second syllable. Stress-assignment errors are rare in nondisordered populations but have been associated with clinical populations (e.g., nonfluent aphasia, apraxia). Nickels and Howard (1999) outline the level of deficit that certain stress-related errors may represent in the context of Levelt's (1989) framework. They posit that errors of stress assignment or omission of unstressed syllables during words with weak–strong (i.e., iambic) stress pattern reflect deficient representations or insufficient retrieval of stored metrical frames. In response, the phonological encoding system reassigns default metrical patterns (i.e., trochaic, or strong–weak) in lieu of access to stored metrical frames. Observed stress-assignment errors in AWS meet these criteria and are perhaps indicative of disrupted retrieval of the nondefault metrical frame during phonological encoding. Findings also introduce the potential impact of nondefault metrical stress on accurate segmental assignment in AWS when planning and producing utterances as short as two syllables, unlike previous studies that found reduced accuracy for longer nonwords (e.g., Byrd et al., 2012; Sasisekaran & Weisberg, 2014).

Working Memory

The increased metrical and segmental errors present for AWS in Experiment 2 and not Experiment 1 as well as the noted theoretical relationship of such errors with working memory abilities also warrant further discussion of the role of short-term working memory in the present study. Previous researchers (Sasisekaran et al., 2006; Wheeldon & Morgan, 2002) acknowledge that silent phoneme monitoring tasks using real-word stimuli impose demands on the phonological working memory system. The nonlexical status and nonorthographic cueing of target stimuli used in the present study undoubtedly heightened these demands. Participants were forced to rely more heavily on their phonological working memory with minimal support from long-term semantic or lexical knowledge. It is possible that the reported reaction time differences may reflect differences in phonological short-term memory rather than online phonological processing.

To minimize the anticipated confound of phonological working memory differences, we provided increased training prior to silent phoneme monitoring tasks within each experimental block (i.e., a minimum of 18 exposures). We also removed manual responses with posttrial errors from latency analyses. Comparison of AWNS performance in the present study to previous studies indicate that the frequency of verbal production errors across positions (Experiment 1: 2.43% phonemic error, 0.66% stress errors; Experiment 2: 3.20% phonemic error, 1.00% stress error) were not disproportionately higher than previous tasks using similar training paradigms with AWNS, who reported fewer than 8% verbal errors across studies (Cholin et al., 2006, 2011; Levelt & Wheeldon, 1994). In addition, when posttrial responses were accurate, our data for AWNS are comparable with the overall latency patterns reported in the literature that used real words with initial stress ($C_1 < C_2 = C_3 = C_4$). Thus, the experimental procedure in the present study was sufficient, at least for AWNS, to capture processes similar to those reported in previous silent phoneme monitoring studies and were not unduly compromised by phonological working memory abilities.

However, given the difficulty AWS demonstrate establishing, maintaining, or retrieving phonological information from working memory (e.g., Bajaj, 2007; Byrd et al., 2012; Jones, Fox, & Jacewicz, 2012), particularly in the absence of semantic information (e.g., Byrd, Sheng, Ratner, & Gkalitsiou, 2015; Sasisekaran & Weisberg, 2014), it is still possible and plausible that our experiments were uniquely challenging to phonological working memory of AWS. Correlational data indicate that performance on forward digit span tasks shared a significant negative correlation with segmental errors for AWS, but not AWNS, across both Experiment 1 and 2 (see Figure 2). These correlations occurred despite similar group performance for AWS during forward digit span tasks in Experiment 1, and superior performance for forward digit span in Experiment 2 (see Table 4). As depicted in Figure 2, group differences for 11 AWS in Experiment 2 can be attributed to a wider variation demonstrated by AWS (range = 8 to 14) as compared to AWNS (range = 8 to 12) in Experiment 2. Of these 11, the three AWS with the highest score (i.e., 14) averaged 0% to 3% false negative errors and 0% to 20% phonemic errors, and the four AWS with lowest scores (i.e., 8 to 9) averaged 20% to 43% false negative errors and 14% to 47% phonemic errors. In contrast, negative correlations were not observed in AWNS; all 11 AWNS scored between 0% to 30% false negative errors and 0% to 20% phonemic errors regardless of digit span performance, further suggesting that weaknesses in phonological working memory uniquely influenced the performance of AWS during experimental tasks.

To account for the potential influence of phonological working memory, forward digit span performance was included as a covariate. In Experiment 1, in which stimuli carried initial stress, AWS did not significantly differ from AWNS in the frequency of segmental (false negative or phonemic errors) or metrical errors (stress assignment errors), suggesting that differences in segmental errors that were present prior to covariate inclusion—that is, within-group differences for AWS, $F(2.17, 43.31) = 3.92$, $p = .024$—may have been mitigated by phonological working memory. This was not the case for Experiment 2. Even after inclusion of forward digit span as a covariate, AWS continued to exhibit significantly greater overt segmental errors (posttrial phonemic errors) and covert segmental errors (false negative errors, particularly at the C_2 position) as well as significantly greater metrical errors (posttrial stress-based errors). These data suggest that although phonological working memory may have contributed to errors in Experiment 2, it did not account for all errors.

Specifically, phonological working memory may partially account for the increased phonemic and stress-based errors observed for AWS in Experiment 2. Morgan, Edwards, and Wheeldon (2013) reported a similar pattern for nonstuttering adults during serial recall of nonwords. In this study, six monosyllabic nonwords were grouped based on stress pattern (i.e., trochaic or iambic). Participants recalled segmental properties of nonwords with trochaic stress patterns with significantly better accuracy than nonwords with iambic stress patterns. Morgan et al. interpreted their findings to suggest that when well-learned language-based properties are present, such as typical metrical stress patterns, these properties are recruited to bolster retention of novel phonological code in phonological working memory. When less common metrical patterns are present, such as noninitial metrical stress (Experiment 2), retention of segmental information in phonological working memory is significantly poorer and commensurate with monotone stress patterns. Thus, recruitment of default metrical encoding due to potential weaknesses in working memory in AWS provide an alternative account for increased phonemic and stress-based errors. However, as noted, controlling for observed weaknesses in phonological working memory did not completely diminish the error differences between groups. This suggests that less common metrical stress not only impacts the retention of nonwords in AWS but may also influence the online syllabification process targeted in this study and that phonological encoding is less efficient in AWS even after consideration of potential weaknesses in nonword retention.

These patterns were exclusive to AWS and only in Experiment 2, suggesting that AWNS with similar and sometimes lower phonological working memory (as measured by performance on forward digit span) were efficient in their ability to accurately encode segmental information regardless of metrical configuration and working memory abilities. Present findings support the body of literature that indicates AWNS have a more robust ability than AWS to establish and retain segmental information (e.g., Byrd et al., 2015; Sasisekaran & Weisberg, 2014) and will inform future research by introducing additional factors, such as metrical configuration, which may contribute to difficulties AWS exhibit maintaining and/or generating word forms during phonological encoding.

Limitations

Several techniques were included in the experimental design to reduce the likelihood of participants developing alternative strategies during manual and posttrial responses. Similar to previous silent phoneme monitoring studies (Wheeldon & Levelt, 1995; Wheeldon & Morgan, 2002), nonorthographic cueing of the target nonword was included to reduce the strategy of "visual scanning" for corresponding orthographic letters. Without the benefit of orthographic input, participants were forced to silently generate phonological segments in nonwords. Even with removal of the orthographic stimuli during experimental tasks, it is possible that the participants opted to generate and retain a visual representation of the nonword between training and the experimental task to increase accuracy. However, this potential confound has been addressed by Wheeldon and colleagues as an unlikely strategy during silent phoneme monitoring tasks in AWNS, given the pattern of the experimental data. For instance, adult participants access graphemic information two to three times more slowly than phonological information when given a verbal prompt (Weber & Castleman, 1970; Weber, Kelley, & Little, 1972). Second, previous data indicate that identification patterns of individual graphemes within letter strings typically result in an M-shaped latency distribution, or W-shaped accuracy distribution, suggesting that AWNS identify word-medial letters faster and more accurately than peripheral graphemes (e.g., Hawelka et al., 2006; Tydgat & Grainger, 2009; Ziegler, Pech-Georgel, Dufau, & Grainger, 2010). Neither pattern emerged in the present study for any participant cohort. Nonetheless, the strategy of visual scanning and perhaps occupying the visuospatial loop within working memory cannot be completely ruled out. However, consistency in performance across studies provides greater confidence that this strategy, if present, was infrequent and had a negligible impact on our data.

Conclusion

Although the present study is preliminary and additional investigations are warranted, our findings suggest that the phonological encoding difficulties previously identified in AWS are not limited to segmental properties. Rather, phonological encoding difficulties extend to both segmental and metrical properties. In the absence of initial stress, AWS exhibit reduced segmental accuracy prior to the syllable boundary and slower encoding of segments immediately following initial syllable boundary assignment. Findings also indicate greater phonemic and stress-based errors when producing novel words with noninitial stress. These patterns were not observed in typically fluent adults. Together, the relationship between metrical and segmental encoding abilities in AWS prior to production may uniquely compromise the ability to establish and/or maintain fluent speech.

Acknowledgments

This project was supported by the ASHFoundation New Century Scholars Doctoral Scholarship awarded to the first author; the National Stuttering Association Canadeo Family Research Award with funding awarded to both authors; the Jesse H. Jones Fellowship awarded to the first author. Data collected at the Dr. Jennifer and Emanuel Bodner Developmental Stuttering Laboratory and the Michael and Tami Lang Stuttering Institute were supported by endowments awarded to the second author. We would like to thank Drs. Thomas Marquardt, Harvey Sussman, and Randy Diehl for their valuable input, Dr. Michael Mahometa for assistance during statistical analyses, and Elizabeth Rives and Chelsea Derouen for data processing and reliability efforts.

References

American Speech-Language-Hearing Association (ASHA). (1997). *Guidelines for audiological screening* [Guidelines]. Retrieved from http://www.asha.org/policy

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

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

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX Lexical Database* [CD]. Philadelphia, PA: University of Pennsylvania, Linguistics Data Consortium.

Bajaj, A. (2007). Working memory involvement in stuttering: Exploring the evidence and research implications. *Journal of Fluency Disorders, 32,* 218–238.

Bosshardt, H.-G. (1993). Differences between stutterers' and nonstutterers' short-term recall and recognition performance. *Journal of Speech and Hearing Research, 36,* 286–293.

Bosshardt, H.-G., & Fransen, H. (1996). Online sentence processing in adults who stutter and adults who do not stutter. *Journal of Speech, Language, and Hearing Research, 39,* 785–797.

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

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

Byrd, C. T., Sheng, L., Ratner, N. B., & Gkalitsiou, Z. (2015). Veridical and false recall in adults who stutter. *Journal of Speech, Language, and Hearing Research, 58,* 28–42. doi:10.1044/2014_JSLHR-S-13-0301

Byrd, C. T., Vallely, M., Anderson, J. D., & Sussman, H. (2012). Nonword repetition and phoneme elision in adults who do and do not stutter. *Journal of Fluency Disorders, 37,* 188–201.

Chen, J.-Y., Chen, T.-M., & Dell, G. S. (2002). Word-form encoding in Mandarin Chinese as assessed by the implicit priming task. *Journal of Memory and Language, 46,* 751–781.

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

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

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

Damian, M. F., & Bowers, J. S. (2003). Effects of orthography on speech production in a form-preparation paradigm. *Journal of Memory and Language, 49,* 119–132.

Demonet, J.-F., Chollet, F., Ramsay, S., Cardebat, D., Nespoulous, J.-L., Wise, R., . . . Frackowiak, R. (1992). The anatomy of phonological and semantic processing in normal subjects. *Brain, 115,* 1753–1768.

Dijkstra, T., Frauenfelder, U. H., & Schreuder, R. (1993). Bidirectional grapheme-phoneme activation in a bimodal detection task. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 931–950.

Gathercole, S. (1995). Is nonword repetition a test of phonological memory or long-term knowledge? It all depends on the nonwords. *Memory and Cognition, 23,* 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: Human Experimental Psychology, 56*(A), 1213–1236.

Hawelka, S., Huber, C., & Wimmer, H. (2006). Impaired visual processing of letter and digit strings in adult dyslexic speakers. *Vision Research, 46,* 718–723.

Hennessey, N. W., Nang, C., & Beilby, J. M. (2008). Speeded verbal responding in adults who stutter: Are there deficits in linguistic encoding? *Journal of Fluency Disorders, 33,* 180–202.

Howell, P. (2011). *Recovery from stuttering*. New York, NY: Psychology Press.

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

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

Jones, R. M., Fox, R. A., & Jacewicz, E. (2012). The effects of concurrent cognitive processing load on phonological processing in adults who stutter. *Journal of Speech, Language, and Hearing Research, 55,* 1862–1875.

Kruyskamp, C. (1961). *Groot woordenboek der nederlandse taal* [Great dictionary of the Dutch language] (8th ed.). Antwerp, Belgium: De Nederlandsche Boekhandel.

Labov, W., Ash, S., & Boberg, C. (2006). *Atlas of North American English: Phonology and phonetics*. Berlin, Germany: Mouton de Gruyter.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

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.

Li, P., Sepanski, S., & Zhao, X. (2006). Language history questionnaire: A web-based interface for bilingual research. *Behavior Research Methods, 38,* 202–210.

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing, 19,* 1–36.

Mason, M. (1982). Recognition time for letters and nonletters: Effects of serial position, array size, and processing order. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 724–738.

Mlinar, R. (2010). *FONRYE English dictionary: Phonetic and syllable search*. Retrieved from http://www.languagebits.com/?=682/

Morgan, J. L., Edwards, S., & Wheeldon, L. R. (2013). The relationship between language production and verbal short-term memory: The role of stress grouping. *The Quarterly Journal of Experimental Psychology, 67,* 220–246.

Natke, U., Grosser, J., Sandrieser, P., & Kalveram, K. T. (2002). The duration component of the stress effect in stuttering. *Journal of Fluency Disorders, 27,* 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,* 190–122.

Newman, R. S., & Ratner, N. B. (2007). The role of selected lexical factors on confrontation naming accuracy, speed, and fluency in adults who do and do not stutter. *Journal of Speech, Language, and Hearing Research, 50,* 196–213.

Nickels, L., & Howard, D. (1999). Effects of lexical stress on aphasic word production. *Clinical Linguistics & Phonetics, 13,* 269–294.

Nippold, M. A. (2012). Stuttering and language ability in children: Questioning the connection. *American Journal of Speech-Language Pathology, 21,* 183–196.

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.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9,* 97–113.

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.

Prins, D., Hubbard, C. P., & Krause, M. (1991). Syllabic stress and the occurrence of stuttering. *Journal of Speech and Hearing Research, 34,* 1011–1016.

Roelofs, A. (2014). Modeling of phonological encoding in spoken word production: From Germanic languages to Mandarin Chinese and Japanese. *Japanese Psychological Research, 57,* 22–37.

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

Sasisekaran, J., & Byrd, C. T. (2013). A preliminary investigation of segmentation and rhyme abilities of children who stutter. *Journal of Fluency Disorders, 38,* 222–234.

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.

Sasisekaran, J., & Weisberg, S. (2014). Practice and retention of nonwords in adults who stutter. *Journal of Fluency Disorders, 41,* 55–71.

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

Smith, A., Sadagopan, N., Walsh, B., & Weber-Fox, C. (2010). Increasing phonological complexity reveals heightened instability in inter-articulatory coordination in adults who stutter. *Journal of Fluency Disorders, 35,* 1–18.

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

Stoel-Gammon, C. (2010). The Word Complexity Measure: Description and application to developmental phonology and disorders. *Clinical Linguistics & Phonetics, 24,* 271–282.

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.

Treiman, R., Fowler, C. A., Gross, J., Berch, D., & Weatherston, S. (1995). Syllable structure or word structure? Evidence for onset rime units with disyllabic and trisyllabic stimuli. *Journal of Memory and Language, 34,* 132–155.

Treiman, R., & Zukowski, A. (1990). Toward an understanding of English syllabification. *Journal of Memory and Language, 29,* 66–85.

Tydgat, I., & Grainger, J. (2009). Serial position effects in the identification of letters, digits, and symbols. *Journal of Experimental Psychology, 46,* 275–280.

U.S. Department of Health and Human Services. (1996). *The international classification of diseases, 9th revision, clinical modification (ICD-9-CM)* (4th ed., Vol 1). Retrieved from Centers for Disease Control and Prevention website: http://www.cdc.gov/nchs/icd/icd9cm.htm

Vincent, I., Grela, B. G., & Gilbert, H. R. (2012). Phonological priming in adults who stutter. *Journal of Fluency Disorders, 37,* 91–105.

Vitevitch, M. S., & Luce, P. A. (2004). A web-based interface to calculate phonotactic probability for words and nonwords in English. *Behavior Research Methods, 36,* 481–487.

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

Weber, R. J., & Castleman, J. (1970). The time it takes to imagine. *Perception and Psychophysics, 8,* 165–168.

Weber, R. J., Kelley, J., & Little, S. (1972). Is visual image sequencing under verbal control? *Journal of Experimental Psychology, 96,* 354–362.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale–Third Edition (WAIS-III)*. San Antonio, TX: Pearson PsychCorp.

Wells, J. (2008). *Longman pronunciation dictionary* (3rd ed.). Essex, England: Pearson Education.

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

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

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

Zhang, J.-J., & Xiao, E.-P. (2008). Phonological encoding in the silent speech of persons who stutter. *Acta Psychologica Sinica, 40,* 263–273.

Zhang, Q., & Damian, M. F. (2009). The time course of segment and tone encoding in Chinese spoken production: An event-related potential study. *Cognitive Neuroscience, 29,* 252–265.

Ziegler, J. C., Pech-Georgel, C. P., Dufau, S., & Grainger, J. (2010). Rapid processing of letters, digits and symbols: What purely visual-attentional deficit in developmental dyslexia? *Developmental Science, 13,* 8–14.

Appendix A

Number of Recruited Participants Who Did Not Meet Inclusionary and Exclusionary Criteria for Experiments 1 and 2

		Inclusionary				Exclusionary			
Experiment	No. recruited	Age, y (18+)	Current medical or speech concerns	Current use of antipsychotic medication	Nonnative English proficiency[a]	Failed hearing[b] or vision screenings[c]	Baseline RT ± 2 *SD*	Removed to balance age and gender of groups	Final cohort
Experiment 1									
AWNS	21	0	2	3	1	2	0	2	11
AWS	16	0	0	3	0	2	0	0	11
Total	37	0	2	6	1	4	0	2	22
Experiment 2									
AWNS	20	0	0	1	1	2	2	3	11
AWS	17	0	2	1	1	1	1	0	11
Total	37	0	2	2	2	3	3	3	22

Note. RT = reaction time; *SD* = standard deviation of the mean of each talker group; AWNS = adults who do not stutter; AWS = adults who stutter. Speech diagnoses included diagnosed or observed articulatory or phonological disturbances other than stuttering.

[a]English proficiency on the basis of a 7-point self-rating scale in Language History Questionnaire (Li, Sepanski, & Zhao, 2006). [b]Binaural pure-tone hearing screening (American Speech-Language-Hearing Association, 1997). [c]Visual acuity screening (U.S. Department of Health and Human Services, 1996).