Published in final edited form as:

J Fluency Disord. 2013 March ; 38(1): 45–58. doi:10.1016/j.jfludis.2012.12.003.

A preliminary investigation of phonological encoding skills in children who stutter

Jayanthi Sasisekaran^{*}, Alison Brady¹, and Jillian Stein¹

Department of Speech-Language-Hearing Sciences, University of Minnesota, United States

Abstract

Purpose—The present study investigated phonological encoding skills in children who stutter (CWS) and those who do not (CNS). Participants were 9 CWS ($M = 11.8$ years, $SD = 1.5$) and 9 age and sex matched CNS ($M = 11.8$ years, $SD = 1.5$).

Method—Participants monitored target phonemes located at syllable onsets and offsets of bisyllabic words. Performance in the phoneme monitoring task was compared to an auditory tone monitoring task.

Results—Repeated measures analysis of the response time data revealed significant $Group \times Task \times Position$ interaction with the CWS becoming progressively slower than the CNS in monitoring subsequent phonemes located within the bisyllabic words; differences were not observed in the auditory tone monitoring task. Repeated measures analysis of the error data indicated that the groups were comparable in the percent errors in phoneme vs. tone monitoring. The CWS group was also significantly slower in a picture naming task compared to the CNS.

Conclusions—Present findings suggest that CWS experience temporal asynchronies in one or more processes leading up to phoneme monitoring. The findings are interpreted within the scope of contemporary theories of stuttering.

Educational objectives—At the end of this activity the reader will be able to: (a) discuss the literature on phonological encoding skills in children who stutter, (b) identify theories of phonological encoding in stuttering, (c) define the process of phonological encoding and its implications for fluent speech, (d) suggest future areas of research in the investigation of phonological encoding abilities in children who stutter.

Keywords

Stuttering; Phoneme monitoring; Phonological encoding

1. Introduction

In speech production, the process of phonological encoding can be defined as “involved in retrieving or building a phonetic or articulatory plan from each lemma or word and the utterance as a whole” (Levelt, 1989). This process serves as an interface between lexical processes and speech motor production (Levelt, 1989; Levelt, Roelofs, & Meyer, 1999) and is critical for incremental speech planning and production.

Multifactorial models of stuttering (Smith & Kelly, 1997; Walden et al., 2012) attribute a critical role to linguistic variables in stuttering causation. For instance, the *Dual Diathesis – Stressor* model of developmental stuttering (Walden et al., 2012) postulated reduced linguistic competence and vulnerabilities in speech-language planning and production as causal variables in childhood stuttering. Several mechanism-specific theories of stuttering have postulated phonological encoding as a causal variable. Howell (2004) in his *EXPLAN* theory proposed that fluency failures occur due to temporal asynchronies between execution (EX) and speech planning (PLAN). He speculated that such asynchronies are caused by difficulties associated with the planning of complex linguistic segments and fast speech rate and the resulting coping strategies adopted by the speaker. Postma and Kolk (1993) proposed the *Covert Repair Hypothesis* (CRH) in which the primary symptoms of stuttering represent overt manifestations of covert corrections of speech plan errors that are caused by delayed phonological encoding of speech sounds. Such theories motivate interest in exploring phonological encoding skills in children who stutter (CWS). In the following section, we will briefly review the main findings from studies of phonological encoding in CWS.

1.1. Phonological encoding skills in children who stutter

There are only a few studies that have directly tested phonological encoding skills in CWS. A primary limitation to designing such studies is the need to identify tasks that tap into the process of phonological encoding. For instance, the priming paradigm offers a viable means of testing the organization and activation of the phonological lexicon and has been used in some studies. Melnick, Conture, and Ohde (2003) investigated phonological encoding in 18, 3 to 5-year-old CWS and an equal number of age-matched children who do not stutter (CNS). They used a priming paradigm with speech reaction times measured from three presentation conditions: no prime, phonologically related prime (initial consonant vowel [CV] or CCV of picture name), and phonologically unrelated prime (different initial CV or CCV). The extent of phonological priming in the related prime condition was comparable for both CWS and CNS. Byrd, Conture, and Ohde (2007) investigated phonological encoding in a picture naming auditory priming paradigm in 26 CWS and 26 CNS. There were 13 3-year-olds and 13 5-year-olds in each talker group. Participants were presented with neutral (tone), holistic, or segmental primes before the onset of target pictures and response time to picture naming was measured from picture onset to the time of initiation of naming. The results revealed that the three year-old CWS and CNS were faster in the holistic priming condition and slower in the segmental priming conditions. The five-year-old CNS were fastest in the segmental condition, but the five-year-old CWS were fastest in the holistic condition. The authors attributed the findings to developmental differences in phonological encoding between the groups. That is, by age five CWS appear to demonstrate a delay in segmental encoding abilities as compared to their typically fluent peers. This finding suggested a potential delay in the transition of phonemic competence from holistic to segmental processing abilities in CWS.

Several indirect sources of evidence support a link between impaired phonological (segmental) encoding and stuttering in CWS. First in this category are studies examining the occurrence of disfluencies in utterances of increasing phonemic complexity. For instance, Wolk, Blomgren, and Smith (2000) studied disfluencies in conversational samples of seven children diagnosed with co-existing stuttering and phonological disorders. They found that the percentage of stuttering in initial consonant clusters with phonological errors was higher than in those without phonological errors. Howell, Au-Yeung, and Sackin (2000) investigated the effect of phonologically complex sounds, such as, late-emerging consonants and consonant clusters, on percentage disfluencies in the conversational samples of 51 children who stutter divided into three age groups (3–11 years, 12–18 years, 18+ years).

Results indicated that the frequency of stuttering was higher on the phonologically complex sounds. The occurrence of stuttering in complex sound combinations, such as, consonant clusters, in the above studies can be attributed to difficulties experienced in encoding such late emerging phonemic units.

Anderson and Byrd (2008) studied the influence of phonotactic probability (the frequency of different sound segments and segment sequences) on disfluencies produced by preschool CWS. Language sample from 19 CWS were analyzed and each stuttered word was randomly paired with a fluent word that closely matched in several grammatical and phonemic level variables. Phonotactic probability values were obtained for the stuttered and fluent words from an online database. The results revealed that although phonotactic probability did not have a significant influence on susceptibility to stuttering, within the stuttered words syllable repetitions were significantly lower in phonotactic probability than other disfluency types. This suggested a role for phonological level processes in the occurrence of syllable repetitions. However, Bernstein Ratner, Newman, and Strekas (2009) investigated the effects of word frequency, neighborhood frequency and density on response time and errors in a confrontation picture naming task in 15 CWS and age and sex-matched CNS between 4; 10 (years; months) and 16; 2. Group differences were not observed for word frequency, neighborhood frequency or density. The authors interpreted the findings to suggest that the organization of the phonological lexicon was comparable in CWS and CNS.

A second source of evidence is from studies that have reported poor performance in nonword repetition. For instance, Hakim and Bernstein Ratner (2004) compared eight CWS (4; 3 to 8; 4, years; months) to age-matched CNS using the Children's Test of Nonword Repetition (CNrep; Gathercole, Willis, Baddeley, & Emslie, 1994). Overall higher percent of phonemic errors were observed in the CWS and significant group differences were observed at the three-syllable level. Anderson, Wagovich, and Hall (2006) compared performance of 12 CWS and age-matched controls between 3 and 5 years of age on the CNrep (Gathercole et al., 1994). CWS exhibited significantly fewer correct productions of two- and three-syllable nonwords and a higher percent of phonemic errors in the three-syllable nonwords compared to the CNS. The fact that more phonemic errors were observed in the CWS group in this task suggests potential difficulties in segmentation and phonological encoding. However, cautious interpretation of these findings is warranted due to other reports of comparable nonword repetition performance. For instance, in the largest study of its kind, Smith, Goffman, Sasisekaran, and Weber-Fox (2012) tested 31 CWS aged 4–5 years in nonword repetition skills and found that performance in this task is determined by language status. About half of the CWS who performed poorly on tests of syntax and phonology had difficulties performing the nonword repetition task while the CWS who did not exhibit associated language difficulties were comparable in behavioral performance to the CNS.

In summary, both direct and indirect sources of evidence support the idea of a phonological encoding deficit in children who stutter. While limited direct sources of evidence exist, the results from such studies have been mixed. Perhaps a reason for the mixed results is that these studies have been conducted in pre-school children between 3 and 5 years of age. Data on spontaneous recovery from stuttering suggest that a majority of these children will eventually recover (Andrews & Harris, 1964; Panelli, McFarlane, & Shipley, 1978; Yairi & Seery, 2010) and therefore, may not necessarily exhibit the same performance profiles in phonological encoding as children who persist in stuttering.

1.2. Purposes of the present study

Several psycholinguistic theories of stuttering attribute a causal role to phonological encoding in stuttering (e.g., the EXPLAN, the CRH). Although several sources of evidence

support altered efficiency in performing phonological encoding in CWS, further exploration of the nature of this link is warranted. For instance, are CWS delayed in the timely encoding of phonemic segments during speech production or do they exhibit more errors in the process compared to CNS or both? *Thus, the primary aim of the present study was to investigate phonological encoding skills in school-age CWS using a novel task that would enable investigation of both the time course and errors in performance.* To this end, a phoneme monitoring during silent picture naming task was employed.

According to Levelt's speech production model, self-monitoring of inner or silent speech occurs at the output of phonological encoding. Levelt (1989) and Levelt et al. (1999) argued that speakers monitor their speech output for errors in the speech plan before sending the code for articulatory planning and execution. Thus, monitoring is a natural sub-process of speech production that requires access to sub-lexical units, such as, phonemes, and can be used as a viable task for studying phonological encoding. Sasisekaran and Weber-Fox (2012) used a silent monitoring task to study the development of phoneme and rhyme monitoring in children between 7 and 13 years of age. They reported shorter response times for rhyme monitoring compared to phoneme monitoring across the different age groups tested. In addition, the speed of performance in phoneme monitoring became faster with age indicating the emergence of cognitive processes such as segmentation skills that are critical to performing the verbal monitoring tasks. In the present study, we use the phoneme monitoring task during silent picture naming to study the process of phonological encoding in older CWS. Based on Wheeldon and Levelt (1995) who tested the time course of phonological encoding, that is, progressive monitoring of phonemes located within syllable onsets and offsets of target words, we assumed that phoneme monitoring within words is reflective of the way phonemes are encoded during speech production.

The phoneme monitoring task during silent naming involves several sub-processes – lexical retrieval, phonological encoding and monitoring of the target phoneme, and motor response. Therefore, to control for the possibility that any observed group differences could be due to processes other than phonological encoding, performance in the phoneme monitoring task was contrasted with performance in three other tasks – a picture naming task, a baseline, auditory tone monitoring task, and a simple motor task. Picture naming involves processes that overlap with phoneme monitoring, including lexical access and encoding, followed by speech planning and execution. The phoneme monitoring task was similar to the auditory tone monitoring task in all aspects with the exception that the former requires lexical access and encoding of individual phonemic segments in order to arrive at a decision. The simple motor task was designed to rule out group differences in simple motor responses, which were an inherent component of the phoneme and auditory tone monitoring tasks.

The following research questions were addressed:

1. Do CWS differ from CNS in of the speed of phoneme monitoring within bisyllabic words?
2. Do CWS differ from CNS in the percent error responses in phoneme monitoring?
3. Do CWS differ from CNS in both the phoneme and auditory tone monitoring tasks or are the differences, if any, specific to any one task?
4. Do CWS differ from CNS in the picture naming and the simple motor tasks?

2. Method

2.1. Participants

Nine CWS (1 female) and nine age- and sex-matched CNS between 10 and 14 years participated in the study (CWS: $M = 11.8$ years, $SD = 1.5$; CNS: $M = 11.8$ years, $SD = 1.5$, $t(16) = 0.04$, $p = 0.48$). All participants were right-handed and native speakers of North American English. Participants from the CWS group were recruited through a summer stuttering camp organized by the Julia Davis Speech-Language-Hearing center and by sending letters to parents in the St. Paul school district. Participants from the CNS group were recruited from a pre-existing database and by word of mouth. The test protocol was administered by the second and third authors under the supervision of the first author. All procedures were approved by the Institutional Review Board, University of Minnesota, and participants received reimbursement for participation.

Based on initial screening all participants had a negative history of: (a) neurological deficits, (b) language, speech, reading, hearing difficulties except stuttering in the CWS group, and (c) current usage of medications likely to affect the outcome of the experiment (e.g., for ADHD and anti-anxiety). All participants passed a hearing screening test performed at 20 dB HL at .5, 1, 2, 4, and 8 kHz in both ears. The parents of all participants reported age and grade-appropriate reading skills.

2.2. Inclusion criteria for children who stutter

Participants in the CWS group met the following criteria:

1. Had received a diagnosis of stuttering by a speech-language pathologist,
2. Were rated by the parent as being 2 or more on a 7-point rating scale of stuttering severity at onset and received a rating between 1 and 7 (minimum rating received – 2, maximum rating received – 6; $M = 3.1$, $SD = 1.4$) at the time of testing, and
3. Were receiving treatment currently.

In the present study inclusion criteria (either as percent stuttered syllables or percent stuttered words) were not set for the CWS to be eligible to participate. However, speech data from a reading sample and spontaneous speech from the clinician–child interaction were collected from all children in the stuttering group and analyzed for disfluencies. Stuttered disfluencies including sound and syllable repetitions, word repetitions (considered as stuttering when the number of iterations were equal to or greater than 3; Yaruss, 1998), prolongations, and blocks were coded from the reading and conversation samples by the second and third authors. The total number of stuttered disfluencies in the entire sample was then counted for each subject. These numbers were then converted to reflect the frequency of disfluencies per 100 syllables. Intra- and inter-judge reliability was calculated on 1/3rd of all samples using the formula $\text{total agreement} / (\text{total agreements} + \text{total disagreements}) \times 100$. An intra-judge reliability score of 94% and inter-judge reliability score of 89% was obtained. The disfluency analysis revealed that on average the CWS group exhibited 5.9% ($SD = 5.2$) disfluencies in conversation and 6.0% ($SD = 10.1$) disfluencies in reading. Analysis of the individual disfluency data revealed that all CWS met the standard criteria of a minimum of 3% stuttered syllables in conversation while four CWS had less than 3% stuttered syllables in reading.

2.3. Vocabulary, short-term memory, articulation and phonemic awareness

A series of tests were administered to evaluate expressive and receptive vocabulary, articulation, short-term memory, and phonemic awareness skills in both groups. Receptive vocabulary was tested using the *Peabody Picture Vocabulary test-Edition IV* (PPVT; Dunn

& Dunn, 1997). Expressive vocabulary was tested using the *Expressive Vocabulary Test* (EVT; Williams, 1997). Short-term memory span was determined using the forward and backward digit span tests (*Wechsler's Memory scale*; Wechsler, 1997). Articulation skills were tested using the Sounds-in-Words section of the Goldman Fristoe Test of Articulation (Goldman & Fristoe, 2000). Segmentation skills were tested using the *Lindamood Auditory Conceptualization Test* (LAC; Lindamood & Lindamood, 1979). The LAC is used to test the cognitive ability to perceive, conceptualize, and manipulate speech sounds; skills that constitute phonemic awareness. Subtest 1 measures participants' familiarity with isolated phoneme and phoneme sequences patterns. For this subtest participants were asked to arrange colored blocks in a sequence depending on how many sounds they heard and the order in which the sounds were repeated within a sequence. Subtest 2 measures phoneme discrimination skills in monosyllables. Participants were asked to re-arrange and add new cubes to a pre-established sequence based on changes to a nonsense syllable sequence. Scores from Subtests 1 and 2 were used to calculate converted scores that enabled comparison at grade level.

2.4. Stimuli, tasks, and procedure

The experiment consisted of four tasks: a simple motor task, a picture familiarization and naming task, a phoneme monitoring task, and an auditory tone monitoring task. The picture familiarization and naming tasks were always presented before the phoneme monitoring task. The order of the remaining three tasks was randomized across participants.

2.4.1. Simple motor task—The stimulus for this task was a 0.5 kHz pure tone of 550 ms duration. The purpose of this task was to study the time taken to execute simple manual responses by participants in both groups. During this task, participants were presented with 26 trials, each consisting of a random inter-stimulus interval blank screen of 700 ms, 1400 ms, 2100 ms followed by a 0.5 kHz pure tone 550 ms in length. Participants were instructed to respond to the onset of the target tone as quickly as possible by clicking the left-most button on the response box.

2.4.2. Picture familiarization and naming—Twelve bisyllabic nouns (CVC(C)CVC, CVCCCV) were the stimuli for the picture naming task (see Appendix A). All target words had the stress placed on the first syllable and carried target consonants at the onset and offset of each syllable. Black and white line-drawings representing the target words were selected from Snodgrass and Vandervart (1980) and used to elicit picture naming responses. Appendix A shows the name agreement (agreement between given name for an object and the name provided), image agreement (how closely each picture resembles the mental image of an object; $M = 3.9$, $SD = 1.0$), word familiarity (judged based on how usual or unusual an object is from experience; $M = 3.6$, $SD = 1.0$), complexity (rating of picture intricacy; $M = 2.6$, $SD = 0.8$), *Kucera–Francis* written word frequency ($M = 14.9$, $SD = 15.3$), and age of acquisition (if available).

The purposes of this task were to familiarize participants with the 12 target pictures and to elicit picture naming responses post familiarization to record response time and errors in picture naming. Participants were presented with pictures corresponding to the target words on a computer screen. They were asked to name each picture and were corrected if any naming errors were observed. This task took a maximum of 5 min for each participant. After familiarization, participants performed the naming task during which each picture was presented twice in random order. Each picture was preceded by a pure tone to indicate picture onset. This was followed by the picture itself and participants were instructed to name each picture as soon as it appeared on the screen.

2.4.3. Phoneme monitoring—The target words and the corresponding pictures from the picture naming task were also used to elicit phoneme monitoring responses. Appendix A shows the four phonemes located at first and second syllable onset and offset positions that were monitored for each word.

The purpose of this task was to measure participants' response time (in ms) and accuracy in monitoring the presence or absence of target phonemes during silent picture naming. Phonemes to be monitored occurred in the first and second syllable onset or offset positions (e.g., $\underline{C}^1VC^2 \underline{C}^3VC^4$ – b¹as²k³et⁴, $\underline{C}^1VC^2 \underline{C}^3C^4V$ – c¹an²d³l⁴e), of the 12 bisyllabic words. The 12 target words were assigned to four blocks such that each target position in a word was monitored once in a block. In each block, the 12 target words occurred twice (once with the phoneme as a target, thus requiring participants to provide a 'yes' response, and once without the phoneme as a target, requiring a 'no' response) for a total of 24 words per block. Within each block, the order of the target words was randomized. The order of the four blocks was counterbalanced across participants.

Participants were seated comfortably in front of a 15-in. computer screen. Prior to the task participants were given the following instructions: *"In this task, you will hear a sound, e.g., /tə/, /pə/, or /fə/, and this will be followed by one of the pictures that you named earlier. You are required to silently name the picture while looking for the presence or absence of the sound in the picture's name. The sound could be present either at the beginning, middle, or end of the picture's name. Press the green button on this box as soon as you identify the target sound in the name and the 'red' button if the sound is absent. You will see the same picture another time after you press the button and this time you have to name the picture aloud. Wait after you name the picture aloud for the next sound and picture."* Following instructions, three to five practice trials were used to familiarize participants with the task.

A trial in each block consisted of the following series of events: (a) an orienting screen for 700 ms followed by auditory presentation of a pre-recorded target phoneme (each target phoneme was always presented along with a schwa vowel although participants were asked to monitor the target phoneme irrespective of the sound preceding or following it); (b) varying inter-stimulus interval (ISI) between hearing the target phoneme and seeing the target picture. ISI was varied to reduce anticipatory button press responses from participants; (c) target picture presented on the screen for 3 s and response time was measured from the onset of the picture. Participants pressing the green ('yes') or the red ('no') button using the index and middle fingers of the dominant hand to indicate the presence/absence of a phoneme in the target; (d) manual response initiating the presentation of the same picture with participants naming the picture aloud. This was done to determine if a child was thinking of the target word, as opposed to another word, when responding to the monitoring task. Presentation of the next trial in the sequence was initiated by the experimenter after participants' response or automatically after 3 s in case of no response.

2.4.4. Auditory tone monitoring—The stimuli for this task were made up of 96 computer-generated auditory stimuli consisting of a sequence of four pure tones, distributed across four blocks. Half of the tone sequences consisted of one 1 kHz tone and three 0.5 kHz tones. The occurrence of the 1 kHz target tone at each of four positions was distributed evenly (e.g., 1 .5 .5 5, .5 1 .5 .5, .5 .5 1 .5, .5 .5 .5 1). The remaining stimuli consisted of four identical 0.5 kHz tones. The overall length of each tone sequence was matched to the average length of the target bisyllabic words in the experiment as produced by a native English speaker and measured using PRAAT. Based on this analysis, the average word duration was determined as 650 ms ($SD = 118$). Therefore, the tone sequences were generated such that the duration of each tone in the sequences was 100 ms with a 50 ms gap between each tone in the sequence.

The purpose of this task was to assess general auditory tone monitoring skills in both groups. The stimuli were made up of 96 computer-generated auditory stimuli consisting of a sequence of four pure tones, distributed across four blocks. Half of the tone sequences in each block ($N = 12$) consisted of one 1 kHz tone and three 0.5 kHz tones. The position of the 1 kHz tone was distributed evenly across the four positions in the four-tone sequence. These tone sequences required a 'yes' response from the participant indicating the presence of the target tone. The other half ($N = 12$) consisted of four .5 kHz tones that constitute the sequences. These tone sequences required a 'no' response from the participants. Within each block, the order of the tone sequences was randomized. The order of the four blocks was counterbalanced across participants.

Prior to the task, participants were familiarized with the tone-sequences and the target 1 kHz tone that they had to monitor across the four positions. Participants were given the following instructions: *"In this task, you will hear a tone first and this will be followed by a four tone-sequence, e.g., beep-beep-beep-beep. You are required to identify if the first tone that you hear is in the four-tone sequence. Press the green button on this box as soon as you identify that the target tone is in the sequence and the 'red' button if not. Wait after you press the button for the next target tone and tone sequence pair."* A trial in each block consisted of the following series of events: (a) an orienting screen for 700 ms followed by auditory presentation of a pre-recorded target 1 kHz tone; (b) inter-stimulus interval (ISI) of 700, 1400 or 2100 ms between hearing the first tone and the subsequent four tone sequence. ISIs were varied to reduce anticipatory button press responses from participants; (c) presentation of the tone-sequence and response time measured from the onset of the tone sequence and participants pressing the green ('yes') or the red ('no') button using the index and middle fingers of the dominant hand to indicate the presence/absence of the target tone in the sequence; (d) manual response initiated the presentation of the next tone-sequence pair.

2.5. Instrumentation

The experimental stimuli were programmed and presented using Super Lab v 4.5 software. A Toshiba laptop was used to present the stimuli for the tasks. Manual responses from the monitoring and simple motor tasks were recorded using the Cedrus response box. Spoken responses from the overt picture naming trials were recorded using a Sony Digital Voice Recorder. Reaction time, the time (in ms) between presentation of the stimuli and subject response across the tasks, was automatically recorded by Super Lab and stored on the laptop's hard drive.

2.6. Data scoring

Trials in each task were categorized as correct, error, and outlier responses. In the monitoring tasks, correct responses included trials where participants identified correctly the presence or absence of a phoneme or tone match as well as named the picture correctly at the end of each phoneme monitoring trial. Correct responses were further analyzed as 'yes' (trials where the target sound was present in the picture's name) and 'no' responses (trials where the target sound was not present in the picture's name). Outlier responses included trials where the response time was 2 *SD* above or below the individual's mean response time for each task and response type. Error responses included both incorrect and absent responses, with incorrect responses including trials where participants responded with a *false positive* or a *false negative* response to the presence or absence of a phoneme or a tone match. Error responses also included trials where the pictures were not named correctly at the end of a phoneme monitoring trial. A total of four trials involving such naming errors were identified and excluded from the data set. Table 1 provides a summary of the percent errors and outliers across the four target positions for both 'yes' and 'no' responses from each group for the phoneme and auditory tone monitoring tasks. In the picture naming task,

a total of 0.27% trials in the CNS and 0.96% trials in the CWS were excluded as errors and/or disfluencies. There were no outliers or errors in the simple motor task. Only correct 'yes' responses across the four target positions were included in the response time analysis while 'no' responses as well as outlier and error responses were excluded. The 'yes' vs. 'no' responses were analyzed separately as were the error responses.

2.7. Statistical analysis

Responses from the simple motor and picture naming tasks were analyzed using independent samples *t*-test. Preliminary analyses of the monitoring time data (in ms) revealed normality of distribution for the phoneme and auditory tone monitoring response time data across the two groups (*Kolmogorov–Smirnov* test, Phoneme monitoring, $p = 0.20$; Tone monitoring, $p = 0.20$). *Levene's* test of homogeneity of variance revealed non-significant p values (>0.05) for all of the dependent variables. The preliminary analyses affirmed the use of parametric statistics for further data analysis. Independent and paired samples *t*-tests were used to analyze the average response times and percent errors from the 'yes' vs. 'no' responses in both the phoneme and auditory tone monitoring tasks across both groups. Two repeated measures analysis were performed on the data from the 'yes' responses. In the first analysis, response time, the primary dependent variable, was analyzed to study differences between the CWS and CNS groups in the two tasks – phoneme and tone monitoring. For this analysis, Group (CWS, CNS) was the between-subjects variable while Task (phoneme, tone) and Position (1–4) were the within-subjects variables. Age was added as a covariate based on earlier reports of an age effect in performing these tasks (Sasisekaran & Weber-Fox, 2012). A second repeated measures analysis was run on the error data to investigate differences between the groups in the percent errors in monitoring across the two tasks. For this analysis, Group (CWS, CNS) was the between-subjects variable while Task (phoneme, tone) and Position (1–4) were the within-subjects variables.

3. Results

3.1. Vocabulary, short-term memory, articulation, and phonemic awareness

Analysis revealed non-significant group differences in tests of receptive (PPVT Standard score: CWS, $M = 113.1$, $SD = 11.2$; CNS, $M = 117.5$, $SD = 16.1$, $t(16) = -67$, $p = 0.25$) and expressive vocabulary (EVT Standard score: CWS, $M = 104.2$, $SD = 13.1$; CNS, $M = 111.4$, $SD = 15.5$, $t(16) = -1.1$, $p = 0.15$). A trend for significance was observed in the test of articulation (GFTA Standard Score: CWS, $M = 102.5$, $SD = 1.2$; CNS, $M = 104.1$, $SD = 2.9$), $t(16) = -1.36$, $p = 0.09$. Non-significant differences were also observed in the forward (CWS, $M = 9.5$, $SD = 1.2$; CNS, $M = 10.7$, $SD = 2.9$), $t(16) = -1.14$, $p = 0.13$, and the backward digit spans (CWS, $M = 7.6$, $SD = 4.5$; CNS, $M = 6.2$, $SD = 3.4$), $t(16) = 0.76$, $p = 0.23$. Comparison of the LAC converted scores revealed significant differences between the groups with the CWS scoring lesser than the CNS, $t(16) = -1.8$, $p = 0.04$ (CWS, $M = 94.4$, $SD = 30.3$; CNS, $M = 115.3$, $SD = 14.6$).

3.2. Simple motor task

An independent samples *t*-test revealed the groups to be comparable in the speed of manual responses (CWS, $M = 165$ ms, $SD = 63.3$; CNS, $M = 176$ ms, $SD = 91.4$), $t(22) = -0.34$, $p = 0.36$.

3.3. Picture naming

An independent samples *t*-test revealed the groups to be significantly different in speed of picture naming, $t(16) = 3.49$, $p = 0.002$. The CWS group ($M = 743.4$ ms, $SD = 106.5$) was significantly slower in the time taken to initiate the picture naming responses compared to the CNS group ($M = 609$ ms, $SD = 43.8$).

3.4. Monitoring time (in ms)

3.4.1. 'Yes' vs. 'no' responses—An independent samples *t*-test of the phoneme monitoring data revealed the groups to be comparable in response times to the 'no' responses, $t(16) = 1.04$, $p = 0.15$ (CWS, $M = 1788.9$ ms, $SD = 515.5$; CNS, $M = 1539$ ms, $SD = 497.1$), while the CWS group ($M = 1789.6$ ms, $SD = 424.5$) was significantly slower than the CNS group ($M = 1421.3$ ms, $SD = 372.9$) in the 'yes' responses, $t(16) = 1.95$, $p = 0.03$. A paired samples *t*-test revealed that the 'yes' responses were comparable in speed to the 'no' responses in both CWS, $t(8) = 0.01$, $p = 0.98$, and CNS, $t(8) = -0.92$, $p = 0.38$. Analysis of the auditory tone monitoring data revealed no significant differences between groups in the 'yes' (CWS, $M = 583.3.9$ ms, $SD = 249.0$; CNS, $M = 515.8$ ms, $SD = 152.8$, $t(16) = 0.69$, $p = 0.24$) or the 'no' responses (CWS, $M = 641.0$ ms, $SD = 301.8$; CNS, $M = 597.4$ ms, $SD = 189.9$, $t(16) = 0.36$, $p = 0.35$). A paired samples *t*-test revealed that the 'yes' responses were comparable in speed to the 'no' responses for the CNS, $t(8) = -1.9$, $p = 0.9$. For the CWS, the 'no' responses were significantly longer than the 'yes' responses, $t(8) = -2.3$, $p = 0.04$.

3.4.2. Monitoring time across target positions—This analysis was done on the 'yes' responses to study further differences between the CWS and CNS groups in the speed of phoneme and tone monitoring across the four target positions within the bisyllabic target words. Group was the between-subjects variable while Tasks (Phoneme, Tone) and Position (1–4) were the within-subjects variable. *Huynh-Feldt* p values are reported for conditions where the sphericity assumption was violated. Including Age as a covariate resulted in a significant main effect of Age, although non-significant interactions of other variables with Age was observed. Furthermore, based on Thomas et al. (2009) that inclusion of a covariate, such as age, results in a weaker main effect of the repeated measure and an overly conservative test of the within subject measures, Age was not retained as a covariate.

The analysis revealed a significant main effect of Group, with the CWS ($M = 1789$ ms, $SD = 424$) being slower than the CNS in performing the monitoring tasks ($M = 1421.3$ ms, $SD = 372$), $F(1, 16) = 4.4$, $p = 0.05$, $\eta^2 = 0.21$. A significant effect of Task, $F(1, 16) = 137.0$, $p = 0.00001$, $\eta^2 = 0.89$ was obtained. This showed that in both groups the auditory tone monitoring task was performed faster ($M = 549$ ms, $SD = 203$) than the phoneme monitoring task ($M = 1605$ ms, $SD = 431$). A significant Position effect, (*Huynh-Feldt*) $F(3, 48) = 37.1$, $p = 0.00001$, $\eta^2 = 0.69$, showed that the monitoring times for target sounds in each position was significantly different from the next position (Position 1, $M = 917$ ms, $SD = 687.4$; Position 2, $M = 1046.5$ ms, $SD = 752.5.4$, Position 3, $M = 1096.2$ ms, $SD = 672.2.2$; Position 4, $M = 1274.0$ ms, $SD = 662.1.4$). A trend for significance was observed in the Group $\times$ Task interaction, (*Huynh-Feldt*) $F(1, 16) = 3.5$, $p = 0.07$, $\eta^2 = 0.17$. A significant Group $\times$ Position effect, (*Huynh-Feldt*) $F(3, 58) = 4.07$, $p = 0.01$, $\eta^2 = 0.20$, indicated significant differences in monitoring times between CWS and CNS for the fourth target position although differences in *Mean* were observed between the groups at all positions. A significant Task $\times$ Position effect, (*Huynh-Feldt*) $F(3, 48) = 26.0$, $p = 0.00001$, $\eta^2 = 0.61$, was also observed. Finally, a significant Group $\times$ Task $\times$ Position interaction, (*Huynh-Feldt*) $F(3, 48) = 4.3$, $p = 0.008$, $\eta^2 = 0.21$, was obtained (see Fig. 1). Post hoc comparisons (*Fisher's LSD*) revealed that the CWS group was significantly slower than the CNS in monitoring target phonemes in the third ($p = 0.01$) and fourth, word-final position ($p = 0.001$). The groups were not significantly different in monitoring times across the first two positions, although the CWS group was slower than the CNS group (Position 1, $p = 0.10$, Position 2, $p = 0.15$). Significant differences were not observed between the groups across the four positions for the auditory tone monitoring task.

3.4.3. LAC converted scores vs. average phoneme monitoring time—Fig. 2 illustrates the LAC scores and average phoneme monitoring times across participants in both groups. Although the correlations were non-significant in both groups (CWS, *Pearson* $r = .07$; CNS, *Pearson* $r = -0.42$), the scatter plot showed that 8 of the 9 CNS participants scored between 100 and 140 in LAC while the CWS group showed higher variability in performance.

3.5. Monitoring errors

3.5.1. 'Yes' vs. 'no' responses—The percent errors in the 'yes' vs. 'no' responses for phoneme and auditory tone monitoring were analyzed separately. An independent samples *t*-test of the phoneme monitoring data revealed the groups to be comparable in the percent errors in both the 'yes', $t(16) = -0.47$, $p = 0.32$ (CWS, $M = 22.7$, $SD = 9.5$; CNS, $M = 24.8$, $SD = 9.2$), and the 'no' responses, $t(16) = -0.81$, $p = 0.21$ (CWS, $M = 6.2$, $SD = 5.9$; CNS, $M = 8.7$, $SD = 7.1$). A paired samples *t*-test revealed that the percent errors in the 'yes' responses were significantly more than in the 'no' responses for both the CWS, $t(8) = 3.7$, $p = .005$, and CNS, $t(8) = 5.16$, $p = 0.0008$. Analysis of the auditory tone monitoring data revealed no significant differences in error rates between groups in the 'yes' (CWS, $M = 5.5$, $SD = 6.1$; CNS, $M = 5.9$, $SD = 4.9$, $t(16) = -0.17$, $p = 0.43$) or the 'no' responses (CWS, $M = 1.8$, $SD = 2.8$; CNS, $M = 3.0$, $SD = 2.9$, $t(16) = -0.84$, $p = 0.20$). A paired samples *t*-test revealed that the percent errors in the 'yes' responses were comparable to the 'no' responses for both the CWS, $t(8) = 1.55$, $p = .15$, and CNS, $t(8) = 1.97$, $p = 0.08$.

3.5.2. Monitoring errors across target positions—This analysis was done to study differences between the CWS and CNS groups in the percent errors in phoneme and tone monitoring. Group was the between-subjects variable while Tasks (Phoneme, Tone) and Position (1–4) were the within-subjects variable. *Huynh-Feldt* p values are reported for conditions where the sphericity assumption was violated. This analysis revealed a non-significant main effect of Group, with the CWS ($M = 22.8$, $SD = 9.5$) and the CNS ($M = 24.9$, $SD = 9.2$) being comparable in the overall percent of monitoring errors, $F(1, 16) = 0.22$, $p = 0.64$, $\eta^2 = 0.01$. A significant effect of Task, $F(1, 16) = 54.8$, $p = 0.00001$, $\eta^2 = 0.77$, was obtained. This analysis revealed a higher percentage of monitoring errors in the phoneme monitoring task ($M = 23.8$, $SD = 9.2$) than in the auditory tone monitoring task ($M = 5.8$, $SD = 5.4$). A significant Position effect, $F(3, 48) = 6.5$, $p = .0008$, $\eta^2 = 0.29$, showed that the syllable offset positions (Positions 2 and 4) had significantly more monitoring errors (Position 1, $M = 9.9$, $SD = 9.7$; Position 2, $M = 18.9$, $SD = 19.9$; Position 3, $M = 12.4$, $SD = 12.5$; Position 4, $M = 17.8$, $SD = 18.9$). A significant Task $\times$ Position effect was also obtained, $F(3, 48) = 12.5$, $p = .0001$, $\eta^2 = 0.43$. Post hoc comparison (*Fisher's LSD*) revealed comparable error rates for the phoneme and auditory tone monitoring tasks only for the word-onset position ($p = 0.53$). The difference in error rate across all other positions for the two tasks was significant. All other interactions were non-significant, Task $\times$ Group, $F(1, 16) = 0.11$, $p = 0.74$; Group $\times$ Position, $F(3, 58) = 0.47$, $p = 0.70$; Group $\times$ Task $\times$ Position, $F(3, 48) = 0.55$, $p = 0.64$ (see Fig. 3).

4. Discussion

The aim of the present study was to investigate phonological encoding abilities in a group of children with chronic stuttering. Earlier studies investigating such skills in younger CWS constitute indirect sources of evidence, while studies that have directly assessed such skills have reported mixed results. This investigation was also motivated by the presence of several psycholinguistic theories of stuttering that attribute a causal role to phonological encoding. Phonological encoding was investigated using a phoneme monitoring in silent naming task. This task enabled an investigation of the encoding of phonemic segments

within words while also affording the opportunity to study error rate in task performance. Performances in the verbal monitoring task was compared to a nonverbal, auditory tone monitoring task to investigate if certain cognitive processes shared across the two tasks may be implicated in CWS. Finally, the groups were compared in a picture naming as well as a simple motor task to determine differences, if any, in manual motor responses.

4.1. Vocabulary, short-term memory, articulation and phonemic awareness

In the present study the groups were compared in measures of receptive and expressive vocabulary, articulation, and short-term memory. Performance in these measures may be relevant to phoneme monitoring performance. For instance, studies in children have shown that vocabulary is a predictor of phonemic competence (e.g., Edwards, Beckman, & Munson, 2004) and thus could potentially influence performance in phoneme monitoring. Poor short-term memory may influence concurrent performance of silent picture naming and phoneme monitoring, tasks that require storage and retrieval of information from short-term storage. The results suggested that the groups were comparable in expressive, receptive vocabulary, and short-term memory while showing a trend for significant group difference in the test of articulation. Of interest was the finding that the CWS scored lower than the CNS in a test of phonemic awareness, namely the LAC. The LAC subtests required participants to identify, segment, and manipulate phonemes in speech, skills that are inherent to the phoneme monitoring task. Lower scores in the LAC as well as a trend for lower scores in the test of articulation suggest that the CWS may have been challenged in such skills. However, present finding of significantly lower LAC scores in CWS is contrary to earlier reports of comparable performance of CWS and CNS in this test. For instance, Bajaj, Hodson, and Schommer-Aikins (2004) tested CWS and CNS between 3 and 5 years in LAC and found no group differences. It is possible that both the younger CWS and CNS tested in Bajaj et al. (2004) were still acquiring segmental processing abilities and were therefore comparable in LAC performance as neither groups were operating at floor or ceiling. Lower group average and higher variability in LAC scores from the present study suggest that some older CWS may continue exhibiting phonemic awareness skills that may not be grade-appropriate.

4.2. Performance in phoneme monitoring and implications

In the present study, the monitoring of target phonemes located in syllable onsets and offsets within bisyllabic words was investigated. The findings revealed that CWS were significantly slower than the CNS in phoneme monitoring. The results from the simple motor task revealed that the observed differences in phoneme monitoring were independent of differences in simple motor responses, which were integral to performing phoneme monitoring. An interpretation is that the CWS are slower in encoding segmental units during speech. This finding supports theories, such as, the EXPLAN (Howell, 2004), that attribute temporal asynchronies in the encoding of phonemic units during speech planning and production to stuttering. Yet another reason for delayed phoneme monitoring in CWS could be reduced proficiency in phoneme awareness skills, such as, segmentation, as supported by poor performance in LAC. Further support for delayed phonological encoding is also evident from slower picture naming responses in CWS. Picture naming involves several processes including, selecting the target word from the mental lexicon, phonological encoding, phonetic encoding, and initiation of articulation (Levelt et al., 1999). A delay in encoding may be one possible explanation for the differences in this task between CWS and CNS. However, other explanations are likely and should be considered in interpreting the above data. For instance, both delayed phoneme monitoring and picture naming responses in CWS can be attributed to difficulties in phonetic encoding leading up to stages of motor execution. Levelt (1989) proposed that the phonetic plan or the gestural score is the representation on which covert monitoring is performed (but see, Jackendoff, 1987, for an

alternate explanation). As further support for this, neuroimaging studies (Ackermann & Riecker, 2004; Yetkin et al., 1995) indicate that covert speech activates many of the motor areas involved in overt speech processes, thereby suggesting that the phonetic plan may indeed be available for covert monitoring. This would suggest that CWS are delayed in the timely generation of a phonetic plan, which is then evident in both the phoneme monitoring and picture naming tasks (but see, van Lieshout, Hulstijn, & Peters, 1996, for evidence to the contrary).

Based on Wheeldon and Levelt (1995), we assumed that the pattern of phoneme monitoring within a word is reflective of the time course of phonological encoding. Present findings revealed that the CWS got progressively slower in monitoring subsequent phonemes within the bisyllabic words. Initially, both the CWS and CNS were comparable in monitoring phonemes located at the onset and offset of the first syllable. Such differences were significant for the phonemes located at the onset and offset of the second syllable. This finding suggested that difficulties experienced by CWS in phonological encoding are likely to be compounded by increasing word length and phonemic complexity. This interpretation is supported by reports of higher percent of stutter events in longer and complex words (e.g., Howell & Au-Yeung, 1995; Soderberg, 1966; Wolk et al., 2000). Comparison of the pattern of monitoring across the four positions within the bisyllabic words revealed that in both groups the phonemes located in the first syllable were monitored faster than those located in the second syllable. This finding confirms other similar reports (e.g., Jansma & Schiller, 2004) and offers support for sequential encoding of speech. Additionally, some differences were also evident in the time course of phoneme monitoring. For instance, evidence for cascaded or parallel encoding of subsequent syllables was evident only in the CNS group. Cascaded encoding is the process by which more than one phoneme may be activated simultaneously within a time frame. Such cascading or parallel encoding may be critical while producing longer utterances under time constraints in spontaneous speech. This was evident in the CNS group as comparable monitoring times for phonemes located in the first syllable offset (*Mean* = 1476 ms, *SD* = 490) and the second syllable onset (*Mean* = 1429 ms, *SD* = 315). A similar pattern of cascaded encoding was not evident in the CWS, that is, the monitoring time for phonemes located in the first syllable offset (*Mean* = 1802 ms, *SD* = 469) was significantly different from the phonemes located at second syllable onset (*Mean* = 1972 ms, *SD* = 492).

The present findings revealed the groups to be comparable in the percent errors in monitoring, although some differences were noted in the percent errors across the different positions. This finding did not support the assertion that the speech plan of CWS may have a higher number of errors (Postma & Kolk, 1993). Instead the findings suggest that any difficulties in encoding are primarily in the time domain. Furthermore, examination of the response time and error data for a speed-accuracy trade-off revealed that slower phoneme monitoring was not accompanied by reduced error rate in CWS. This finding suggested that the CWS did not exhibit the typical speed-accuracy trade off strategies that are encountered in response time tasks. However, it could be argued that comparable errors rates between groups in the face of slower monitoring in CWS may be the result of post-lexical search strategies adapted to compensate for such errors, which may have resulted in slower monitoring times. Therefore, further testing using other paradigms is required to confirm present findings of slower phonological encoding in CWS.

4.2.1. “Yes” vs. “no” responses—In the present study, CWS were significantly slower than CNS in the ‘yes’ responses to phoneme monitoring, but not in the ‘no’ responses. Coltheart, Rastle, Perry, Ziegler, and Langdon (2001), Grainger and Jacobs (1996) proposed that ‘yes’ and ‘no’ responses are based on inherently different underlying mechanisms. In lexical decision tasks, ‘yes’ responses are based on unique word identification and global

activation (intra-lexical criteria) while ‘no’ responses are made when the subject has not made a “yes” response before reaching a time limit (extra-lexical criterion). Further support for phonological encoding difficulties in CWS was seen in the significantly slower monitoring times to the ‘yes’ responses. This suggests that the CWS were delayed in meeting the criteria used for making phoneme monitoring decisions, such as, achieving above threshold activation of the target phoneme in order to make a ‘yes’ decision. Similar differences were not observed in the ‘no’ responses and suggests that perhaps CWS may have also been performing additional post-lexical search strategies to confirm the response once they encounter a target phoneme. The findings also revealed that the ‘yes’ and ‘no’ responses were comparable within groups. One potential limitation of the present design that may have resulted in this finding is that each word was presented twice in a block of stimuli; once with and once without the target phoneme. That is, if a subject had encountered a word once and provided a response, he/she could predict the response on the second encounter. The ‘yes’ responses would have been minimally affected by this prediction bias as the location of the target phoneme was distributed across four target positions. The ‘no’ responses would have been affected to a greater extent. Thus, it is likely that the response time of the ‘no’ responses was shortened due to this bias thereby minimizing potential differences between the ‘yes’ and the ‘no’ responses. Both groups also exhibited a higher percent of errors in the ‘yes’ responses compared to the ‘no’ responses. This further supported the earlier assertion that the CWS did not exhibit a speed-accuracy trade-off although the CNS showed faster monitoring speed and therefore more errors in the ‘yes’ responses.

4.2.2. Phoneme vs. auditory tone monitoring—Present findings show that any observed difficulties in monitoring observed in CWS are limited to the verbal monitoring task. No such differences were observed between CWS and CNS in an auditory tone monitoring task similar in design to the phoneme monitoring task and requiring participants to monitor the presence or absence of a target tone with four target positions of a tone sequence. Comparable performance in the auditory tone monitoring task suggested comparable monitoring skills between CWS and CNS. This finding also rules out potential interpretation of the slower phoneme monitoring observed in CWS as reflective of hyper-monitoring (e.g., Bernstein Ratner, 1997) rather than slower phonological encoding, although one could argue that the auditory tone monitoring task in this study did not involve speech stimuli and difficulties in monitoring in CWS may be limited to speech per se. Therefore, further testing is required in order to ascertain if CWS exhibit altered speech monitoring skills independent of phonological encoding abilities.

4.3. Conclusions and future directions

The findings of slower phoneme monitoring in the absence of similar difficulties in auditory tone monitoring suggest a delay in the encoding of phonemes in CWS. This interpretation was further supported by group differences specifically to the ‘yes’ responses, which suggested a delay in meeting the intra-lexical criteria used in making phoneme monitoring decisions in CWS. The findings also suggest that such a delay in phonological encoding is likely to compound with increasing length and phonemic complexity of words. Findings from the present study indicate that reduced phoneme competence, as indicated by performance in LAC, and delayed phonological encoding, as indicated by slower phoneme monitoring performance, could be potential causal variables of stuttering. A limitation of the findings is that stuttering occurs in overt speech while the present findings are from a covert speech task. Therefore, further experimental studies are required to understand the mechanism underlying the transition of the observed delay in phonological encoding to stuttering in overt speech.

Acknowledgments

This study was funded by an NIH R03 grant (R03 DC010047) awarded to the first author. We thank our participants, acknowledge Linda Hinderschiet for assistance with testing, Dr. Edward Carney for technical assistance. We also thank the St. Paul Public School Board for assistance with subject recruitment.

References

- Ackermann H, Riecker A. The contribution of the insula to motor aspects of speech production: A review and a hypothesis. *Brain and Language*. 2004; 89(2):320–328. [PubMed: 15068914]
- Anderson JD, Byrd CT. Phonotactic probability effects in children who stutter. *Journal of Speech, Language, and Hearing Research*. 2008; 51:851–866.
- Anderson JD, Wagovich SA, Hall NE. Nonword repetition skills in young children who do and do not stutter. *Journal of Fluency Disorders*. 2006; 31:177–199. [PubMed: 16814376]
- Andrews, G.; Harris, M. *Clinics in Developmental medicine*. London: Heinemann; 1964. The syndrome of stuttering.
- Bajaj B, Hodson B, Schommer-Aikins M. Performance on phonological and grammatical awareness metalinguistic tasks by children who stutter and their fluent peers. *Journal of Fluency Disorders*. 2004; 29(1):63–77. [PubMed: 15026215]
- Bernstein Ratner, N. Stuttering: A psycholinguistic perspective. In: Curlee, R.; Siegel, GM., editors. *Nature and treatment of stuttering: New directions*. 2. Boston, MA: Allyn & Bacon; 1997. p. 99–127.
- Bernstein Ratner N, Newman R, Streckas A. Effects of word frequency and phonological neighborhood characteristics on confrontation naming in children who stutter and normally fluent peers. *Journal of Fluency Disorders*. 2009; 34(4):225–241. [PubMed: 20113768]
- Byrd CT, Conture EG, Ohde RN. Phonological priming in young children's picture naming: Holistic versus incremental processing. *American Journal of Speech, Language, Pathology*. 2007; 16(1):43–53. [PubMed: 17329674]
- Coltheart M, Rastle K, Perry C, Ziegler J, Langdon R. DRC: A Dual Route Cascaded model of visual word recognition and reading aloud. *Psychological Review*. 2001; 108:204–256. [PubMed: 11212628]
- Dunn, L.; Dunn, L. *Peabody Picture Vocabulary Test-Third Edition*. Circle Pines, MN: American Guidance Service; 1997.
- Edwards J, Beckman ME, Munson B. The interaction between vocabulary size and phonotactic probability effects on children's production accuracy and fluency in nonword repetition. *Journal of Speech, Language, and Hearing Research*. 2004; 47:421–436.
- Gathercole S, Willis C, Baddeley A, Emslie H. The Children's Test of Nonword Repetition: A test of phonological working memory. *Memory*. 1994; 2:103–127. [PubMed: 7584287]
- Goldman, R.; Fristoe, M. *Goldman-Fristoe Test of Articulation*. 2. San Antonio, TX: Pearson Assessments; 2000.
- Grainger J, Jacobs AM. Orthographic processing in visual word recognition: A multiple readout model. *Psychological Review*. 1996; 103:518–565. [PubMed: 8759046]
- Hakim HB, Bernstein Ratner N. Nonword repetition abilities of children who stutter: An exploratory study. *Journal of Fluency Disorders*. 2004; 29:179–199. [PubMed: 15458830]
- Howell P. Assessment of some contemporary theories of stuttering that apply to spontaneous speech. *Contemporary Issues in Communication Sciences and Disorders*. 2004; 31:123–141.
- Howell P, Au-Yeung J. Syntactic determinants of stuttering in the spontaneous speech of normally fluent and stuttering children. *Journal of Fluency Disorders*. 1995; 20:317–330.
- Howell P, Au-Yeung J, Sackin S. Internal structure of content words leading to lifespan differences in phonological difficulty in stuttering. *Journal of Fluency Disorders*. 2000; 25(1):1–20. [PubMed: 18259599]
- Jackendoff, R. *Consciousness and the computational mind*. Cambridge, MA: MIT Press; 1987.
- Jansma BM, Schiller NO. Monitoring syllable boundaries during speech production. *Brain and Language*. 2004; 90:311–317. [PubMed: 15172548]

- Levelt, WJM. Speaking: From intention to articulation. Cambridge, MA: MIT Press; 1989.
- Levelt WJM, Roeloffs A, Meyer AS. A theory of lexical access in speech production. *Behavioral and Brain Sciences*. 1999; 22:1–75. [PubMed: 11301520]
- Lindamood, CH.; Lindamood, PC. The Lindamood Auditory Conceptualization Test (LAC)-Revised edition. DLM Teaching Resources, Allen; Texas: 1979.
- Melnick KS, Conture EG, Ohde RN. Phonological priming in picture naming of young children who stutter. *Journal of Speech, Language, and Hearing Research*. 2003; 46(6):1428–1444.
- Panelli, McFarlane; Shipley. Implications of evaluating and intervening with incipient stutterers. *Journal of Fluency Disorders*. 1978; 3:41–50.
- Postma A, Kolk H. The Covert Repair Hypothesis: Prearticulatory repair processes in normal and stuttered disfluencies. *Journal of Speech and Hearing Research*. 1993; 36(3):472–487. [PubMed: 8331905]
- Sasisekaran J, Weber-Fox C. Cross-sectional study of phoneme and rhyme monitoring abilities in children between 7 and 13 years. *Applied Psycholinguistics*. 2012; 33:253–279. [PubMed: 23204597]
- Smith A, Goffman L, Sasisekaran J, Weber-Fox C. Language and motor abilities of preschool children who stutter: Evidence from behavioral and kinematic indices of nonword repetition performance. *Journal of Fluency Disorders*. 2012; 37(4):344–358. [PubMed: 23218217]
- Smith, A.; Kelly, E. Stuttering: A dynamic, multifactorial model. In: Curlee, RF.; Siegel, GM., editors. *Nature and treatment of stuttering: New directions*, Needham Heights. Vol. 2. MA: Allyn & Bacon; 1997. p. 204-217.
- Snodgrass JG, Vandervart M. A standardized set of 260 pictures: Names for name agreement, familiarity and visual complexity. *Journal of Experimental Psychology*. 1980; 6:174–215.
- Soderberg GA. The relations of stuttering to word length and word frequency. *Journal of Speech and Hearing Research*. 1966; 9:584–589.
- Thomas MSC, Annaz D, Ansari D, Serif G, Jarrold C, Karmiloff-Smith A. Using developmental trajectories to understand developmental disorders. *Journal of Speech, Language, and Hearing Research*. 2009; 52:336–358.
- van Lieshout PHHM, Hulstijn W, Peters HFM. Speech production in people who stutter: Testing the motor plan assembly hypothesis. *Journal of Speech and Hearing Research*. 1996; 39:76–92. [PubMed: 8820700]
- Walden TA, Frankel CB, Buhr AP, Johnson KN, Couture EG, Karrass JM. Dual Diathesis-Stressor Model of Emotional and Linguistic Contributions to Developmental Stuttering. *Journal of Abnormal Child Psychology*. 2012; 4:633–644. [PubMed: 22016200]
- Wechsler, D. WAIS-III administration and scoring manual. San Antonio, TX: The Psychological Corporation; 1997.
- Wheeldon LR, Levelt WJM. Monitoring the time course of phonological encoding. *Journal of Memory and Language*. 1995; 34:311–334.
- Williams, KT. EVT-2: Expressive Vocabulary Test. 2. Bloomington, MN: Pearson; 1997.
- Wolk L, Blomgren M, Smith AB. The frequency of simultaneous disfluency and phonological errors in children: A preliminary investigation. *Journal of Fluency Disorders*. 2000; 25(4):269–281.
- Yairi, E.; Seery, EH. *Stuttering: Foundations and Clinical Applications*. New Jersey: Pearson Education, Inc; 2010.
- Yaruss JS. Real-time analysis of speech fluency: Procedures and reliability training. *American Journal of Speech-Language Pathology*. 1998; 7(2):25–37.
- Yetkin ZF, Hammeke AT, Swanson SJ, Morris GL, Mueller WM, McAuliffe TL, et al. A comparison of functional MR activation patterns during silent and audible language tasks. *American Journal of Neuro-radiology*. 1995; 16:1087–1092.

Appendix A

Snodgrass and Vandervart measure of name agreement (percentage agreement), the *M* and *SD* for ratings of image agreement, familiarity and complexity; the Kucera–Francis written

frequency counts (K–F) for each single-word name; and the Carroll–White age-of-acquisition norms (A–A) for the concepts for which they were available.

	Target phonemes	Name agree. %	<u>Image agree</u>		<u>Familiarity</u>		<u>Complexity</u>		K–F	A–A
			<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
Basket	/b//s//k//t/	90	2.62	1.23	2.18	0.97	4.3	0.84	17	3.12
Candle	/k//n//d//l/	100	3.85	0.76	3.08	1.15	2.4	0.9	18	–
Doorknob	/d//r//n//b/	90	3.9	1	4.25	0.92	2.68	0.61	3	–
Finger	/f//ŋ//g//r/	71	4.6	0.66	4.78	0.79	2.3	0.95	40	–
Football	/f//t//b//l/	100	4.18	0.92	3.55	1.24	2.28	0.71	36	–
Hanger	/h//ŋ//g//r/	86	4.73	0.55	4.52	0.67	1.2	0.56	0	–
Sandal	/s//n//d//l/	Not available in Snodgrass and Vandervart								
Necklace	/n//k//l//s/	60	3.32	1.49	2.7	1.31	1.78	0.88	3	–
Pencil	/p//n//s//l/	100	4.4	0.8	4.42	1	2.32	0.91	34	–
Pumpkin	/p//m//k//n/	98	4.18	1.18	3.08	1.35	2.7	0.7	2	3.69
Sailboat	/s//l//b//t/	93	3.25	0.99	2.92	1.17	3.58	0.92	1	–
Sandwich	/s//n//w//tʃ/	100	3.55	0.97	4.45	0.97	3.42	0.86	10	–

Fig. 1.
Average monitoring times by position and task for the CWS and CNS groups.

Fig. 2.
Scatter plot of average LAC converted scores and phoneme monitoring time in both groups.

Fig. 3.
Average monitoring errors (%) by position and task for the CWS and CNS groups.

Table 1

Percent errors and outliers (and *SD*) for 'yes' vs. 'no' responses by group, task, and position.

Phoneme monitoring											
"Yes" responses				"No" responses							
	Position 1	Position 2	Position 3	Position 4	Position 1	Position 2	Position 3	Position 4			
% errors	CWS	<i>M</i>	11.1	33.3	14.7	32.0	0.0	6.5	6.5	12.0	
		<i>SD</i>	7.2	17.2	14.3	18.2	0.0	9.1	9.1	11.9	
	CNS	<i>M</i>	11.1	36.1	23.1	29.3	10.2	11.1	9.2	4.6	
		<i>SD</i>	7.2	15.6	13.0	20.2	13.0	11.0	13.4	4.4	
% outliers	CWS	<i>M</i>	6.5	1.8	2.8	0.9	5.5	3.7	5.5	5.5	
		<i>SD</i>	3.7	3.7	4.2	2.8	4.2	4.4	4.2	4.2	
	CNS	<i>M</i>	5.1	2.8	3.7	1.8	6.5	4.6	2.8	2.8	
		<i>SD</i>	4.0	4.2	4.4	3.7	3.7	4.4	4.2	4.2	

Auditory monitoring											
"Yes" responses				"No" responses							
	Position 1	Position 2	Position 3	Position 4	Position 1	Position 2	Position 3	Position 4			
% errors	CWS	<i>M</i>	9.2	3.7	5.5	3.7	0.0	2.8	1.8	2.8	
		<i>SD</i>	12.8	6.0	5.9	6.0	0.0	5.9	3.7	8.3	
	CNS	<i>M</i>	8.3	2.8	6.5	6.5	2.8	1.8	0.9	6.5	
		<i>SD</i>	11.8	5.9	6.9	6.9	5.9	5.5	2.8	10.8	
% outliers	CWS	<i>M</i>	5.5	4.6	4.6	6.2	4.6	5.5	2.8	7.1	
		<i>SD</i>	4.2	4.4	4.4	3.5	4.4	4.2	4.2	2.7	
	CNS	<i>M</i>	3.7	6.4	2.8	6.2	4.6	6.5	4.5	5.3	
		<i>SD</i>	4.4	5.5	4.2	3.5	6.0	3.7	4.3	4.0	