Published in final edited form as:

J Fluency Disord. 2014 September ; 41: 55–71. doi:10.1016/j.jfludis.2014.02.004.

Practice and retention of nonwords in adults who stutter

Jayanthi Sasisekaran, Ph.D. and

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

Sanford Weisberg, Ph.D.

School of Statistics, University of Minnesota

Abstract

Purpose—We investigated short-term practice and retention of nonwords in 10 adults who stutter (*Mean* age = 30.7 years, *SD* = 15.1) and age and sex-matched 10 control participants (*Mean* age = 30.8 years, *SD* = 14.9).

Methods—Participants were required to repeat nonwords varying in length (3, 4, 6 syllables), phonotactic constraint (PC vs. NPC, on 3-syllable nonwords only), and complexity (simple, complex). They were tested twice with one hour gap between sessions.

Results—Logistic mixed model of speech accuracy revealed that the AWS showed a significantly lower probability of correct responses with increasing length and complexity. Analysis of speech kinematics revealed practice effects within Session 1 in AWS seen as a reduction in movement variability for the 3-syllable nonwords; the control group was performing at ceiling at this length. For the 4-syllable nonwords, the control group showed a significant reduction in movement variability with practice, and retained this reduction in Session 2, while the AWS group did not show practice or retention. Group differences were not evident at the 6-syllable level.

Conclusions—Group differences in speech accuracy suggest differences in phonemic encoding and/or speech motor processes. Group differences in changes in movement variability within and between sessions suggest reduced practice and retention in AWS. Relevance of the combined use of both behavioral and kinematic measures to interpret the nature of the skill acquisition deficit in persons who stutter is discussed.

Keywords

Stuttering; speech skill acquisition; speech accuracy; movement variability

Stuttering is a developmental speech disorder characterized by the presence of repetition, prolongations, and blocks in speech. Several theories have been proposed to account for stuttering as a deficit in sensorimotor processes (Kalvaram, 2007; Max et al., 2004; Neilson & Neilson, 1991; Smith & Kelly, 1997). The investigation of such processes has been undertaken using both speech and non-speech tasks and the findings allude to a deficit in skill acquisition (Bauerly & De Nil, 2011; Namasivayam & van Lieshout, 2008; Smith, Sadagopan, Walsh, Weber-Fox, 2010; Smits-Bandstra, De Nil, Saint-Cyr, 2006), although some mixed findings have been reported. The aim of the present study is to investigate skill acquisition abilities in the short-term in adults who stutter (AWS) using nonword repetition,

which requires the encoding, assembly, and execution of previously known or unknown speech sound combinations and the constituent motor gestures.

Motor skill acquisition: Theory

Motor skill acquisition refers to the process by which movements produced alone or in a sequence come to be performed effortlessly through repeated practice and interactions with the environment (Willingham, 1998). Of the different stages involved in skill acquisition, studies in the speech motor literature have focused on *practice* (defined as improvements in performance within a single session) and *retention* (defined as performance of the practiced skill between sessions without any practice in the intervening duration between sessions) in typical and atypical (e.g., stuttering) populations (e.g., Behrman, Cauraugh, & Light, 2000; Namasivayam & van Lieshout, 2008; Smith et al., 2010; Smits-Bandstra et al., 2006).

A theoretical notion that has received some attention in the motor literature is the *neuromotor noise hypothesis* according to which changes with motor learning are the consequences of reducing variability in neural command signals (Kleim et al., 2002, 2003; Newell, Liu, Mayer-kress, 2003). Studies of speech and non-speech motor learning suggest that both children and adults exhibit higher levels of neural noise during the acquisition of new movement sequences, evident as higher movement variability (Green, Moore, & Riley, 2002; Smith & Zelaznik, 2004; Takahashi et al., 2002; Vangalen, Portier, Smitsengelsman, & Schomaker, 1993; Walsh, Smith, & Weber-Fox, 2006; Yan, Thomas, Stelmach, & Thomas, 2000). With maturation and practice, increasing neuronal synchronization results in reduced neuromotor noise levels, thereby facilitating motor coordination through the formation of optimal movement synergies. Thus, movement coordination is achieved through a neurophysiologically mediated motor learning process (Harris & Wolpert, 1998; Smith & Zelaznik, 2004).

Skill Acquisition in AWS

Stuttering has been identified as a form of ‘speech clumsiness’ with limitations in acquiring speech motor skills (van Lieshout, Hulstijn, & Peters, 2004). Studies in the stuttering literature have reported reduced practice and longer-term learning (e.g., retention) effects in both non-speech and speech tasks (Adams & Hayden, 1976; Cross & Luper, 1979; Cooper & Allen, 1977; Ludlow et al., 1997; Namasivayam & van Lieshout, 2008; Smith et al., 2010; Smits-Bandstra et al., 2006). The focus of this review is on such studies using speech tasks.

Van Lieshout, Hulstijn, and Peters (1996) tested 12 AWS and 12 control participants in speech reaction time and word duration in a picture and a word naming task. Target words were 1-, 2-, or 3-syllables in length (word size). They reported that AWS were slower in speech reaction times than controls, but the hypothesized group x word size effect was not found. The authors interpreted this finding to suggest that the AWS and controls do not differ in the assembly of the motor plan during speech production. They also reported that word duration was longer in the AWS group and group differences were larger with increasing word size. For both these findings, the authors also reported delays in relative timing of motor events (e.g., onset of upper lip muscle contraction *vs.* thoracic compression). The findings were interpreted to suggest that AWS may use different motor

control strategies to compensate for reduced verbal motor skill in the form of a sensori-motor integration deficit.

Smits-Bandstra et al. (2006) compared nine AWS and nine matched control participants on speed and accuracy in a speech (repeated reading of a 10-syllable nonsense word) and a non-speech (10-item finger tapping) sequence skill learning task. Participants were tested during a practice phase (30 trials), then a transfer phase where two novel sequences were tested. This was followed by retesting of the original sequence used in practice after a 40-minute retention phase. In the finger tapping task, the controls showed practice and retention effects seen as an improvement in finger sequence duration and reaction time while the AWS did not show similar improvements. Group differences with practice were, however, not evident for the speech task either for sequence duration or reaction time. The groups were also comparable in accuracy for both the speech and non-speech tasks. Perhaps, deficits in skill learning may become evident in speech tasks when task complexity is manipulated to reflect a hierarchy of task demands. In this case, the use of a relatively simple nonword sequence (e.g., /ga ba pa ta pa ba ga pa ta ga/) may not have been sufficient to reveal subtle difficulties in skill acquisition in AWS.

Namasivayam and van Lieshout (2008) tested practice and longer-term learning (consolidation, retention) effects on the stability of movement kinematics and the strength and stability of inter-gestural coordination patterns of the upper lip – lower lip – jaw synergy in five AWS and five control participants. Participants produced the bisyllabic nonwords /bapi/ and /bipa/ approximately 15 times across three test sessions spanning a week (T1, T2 in a day, T3 a week apart). Comparison of practice effects (T1 vs. T2) revealed that the control participants showed higher coherence and descriptively lesser variability in gestural coordination thereby showing larger gains with practice. Comparison of learning (retention) effects (T1 vs. T3) revealed that the control group showed higher mean coherence values in T3 compared to T1 while the AWS did not show similar changes.

Bauerly and De Nil (2011) studied practice and retention in 12 AWS and 12 controls. Participants were required to repeat a 11-syllable nonword 100 times (divided into 10 blocks) on Day 1 and 50 times (divided into 5 blocks) on Day 2 session conducted after 24 hours. They measured accuracy, response preparation time, and sequence duration on Day 1 as a measure of practice and on Day 2 as a measure of motor learning (retention). Results failed to confirm the hypothesized poor practice and learning effects in AWS as the expected interaction of group and practice was not evident for accuracy, response anticipation time, or sequence duration on Day 1 or Day 2. The AWS did show significantly slower sequence duration both during practice as well as post-consolidation, which was interpreted as a motor skill limitation. The authors speculated that the relatively high task demands may have resulted in the observed group differences in the speech task contrary to the lack of such differences in Smits-Bandstra et al. (2006).

Smith et al. (2010) systematically varied nonword length (1 to 4 syllables) and complexity (at the 4-syllable level) to study the effects of practice on speech movements in a nonword repetition task in 17 AWS and 17 controls. Participants were required to repeat four nonwords varying in length from 1 to 4 syllables and a fifth, simpler nonword added as a

length control to the 4 syllable nonword. Speech accuracy, lip aperture variability (LA VAR) - a measure of inter-articulator coordination, and speech duration were measured and differences in the speech kinematic measures from the early vs. later five trials were reported. The findings revealed overall higher movement variability and increasing movement variability with increasing nonword length and complexity in AWS. The AWS also showed practice effects in movement coordination seen as an increase in coordination from the early to the later trials within the session while the controls seemed to be operating at ceiling levels of coordination consistent throughout the ten trials. Group differences were not evident in speech accuracy.

Although the above-discussed studies also reported speech errors, the primary focus was on speech motor output as measured by speech reaction time and/or speech kinematics. However, nonword repetition taps into several other underlying processes, including phonological encoding and working memory capacities (Gathercole, 2006), that may potentially improve with practice and retention. Examining speech errors with practice and learning can elucidate the role of these processes in speech skill acquisition in AWS. For instance, using lengthy nonword strings repeated several times, Ludlow et al. (1997) in a preliminary study tested five AWS and five typically fluent adults in their ability to repeat the nonwords and the changes in the percent consonants correct with practice. They found that the AWS were slower than the controls in demonstrating improved accuracy with repeated practice and interpreted the findings to suggest impaired speech and language processes. Byrd, Vallety, Anderson, and Sussman (2012) used a nonword repetition task to investigate phonological working memory skills of AWS. In addition, they reported the number of attempts needed to produce the nonword correctly, an indirect measure of practice. Participants were 14 AWS and 14 age and gender matched controls. Participants had to repeat a set of 12 low phonotactic probability nonwords across four syllable lengths (2, 3, 4, and 7 syllables). AWS were significantly less accurate than the control participants in their initial attempts to produce the longest nonwords (i.e., 7 syllables). The groups were comparable in performance at the other syllable lengths. The AWS also required a significantly higher mean number of attempts to accurately produce 7-syllable nonwords than the controls. The authors interpreted the group differences at the 7-syllable level to suggest phonological working memory deficits in AWS. Such findings carry potential implications for the hypothesized skill acquisition deficit in AWS to possibly involve multiple sources, including phonological encoding, working memory, and/or speech motor output processes.

Purposes of the present study

According to Adam (1987), any skilled behavior involves the combined influences of cognitive, perceptual, and motor processes with different weights. Studies employing kinematic measures to investigate skill acquisition in AWS have reported reduced stability of inter-articulatory coordination with repeated practice (e.g., Smith et al., 2010) and/or motor learning (e.g., Namasivayam & van Lieshout, 2008). However, such studies vary considerably on several factors, including task complexity, extent of practice, and the time to retention. Studies presenting speech error data have reported mixed results (e.g., Byrd, et al., 2012; Ludlow et al., 1997; cf. Bauerly & De Nil, 2011; Smith et al., 2010; Smits-

Bandstra et al., 2006). Those studies reporting comparable group performance (e.g., Bauerly & De Nil, 2011; Smits-Bandstra et al., 2006) have primarily tested practice and learning using fixed syllable sequences consisting of simple (e.g., Consonant-Vowel) syllabic structures. Systematically varying nonword properties, such as, syllable length, phonotactic constraint (PC, *defined* as permissible syllable structure and sound combinations in English), and complexity (operationally *defined* in this study based on syllable-internal phonemic composition and the number, nature of consonant clusters, see ‘Stimuli’ for more details), and studying the effects of these manipulations on speech accuracy and speech kinematics offers an opportunity to identify subtle difficulties experienced by AWS in the different processes potentially contributing to speech skill acquisition. Therefore, our first aim was to investigate the effects of nonword type (this term refers to the two factors, length and PC) and complexity on nonword repetition performance in AWS and controls. Based on the mixed findings reported in the literature (e.g., Byrd et al., 2012; Ludlow et al., 1997; cf. Smith et al., 2010), we hypothesized that behavioral performance in AWS will be influenced by nonword type and complexity and reveal difficulties, if any, in phonemic encoding/working memory and/or speech motor processes.

Learning to produce nonwords requires activation and encoding of phonemic segments that constitute the nonwords and assembling previously known or unknown (in case of nonwords consisting of phonemic combinations that do not follow English PC) motor gestures affiliated with the segments (e.g., Nam, 2007; Tilsen & Goldstein, 2012). Furthermore, retention requires holding the learned information, that is, phonemic spell-out and motor sequencing, in memory for longer durations. The second aim of this study was to investigate within- and between-session changes in behavioral and kinematic measures to test practice and retention in the nonword repetition task in AWS and controls. The effects of practice on speech errors and speech kinematics were tested in Session 1. Retention was tested an hour after the first session, without any latent practice during the intervening hour. Lip aperture variability index (LA VAR; Smith, Johnson, McGillem, & Goffman, 2000), a composite measure of spatial and temporal movement variability of the lip aperture trajectory for repeated productions, was the kinematic measure used to investigate changes in coordinative stability of the upper lip – lower – jaw synergy with practice and retention. Smith and Zelaznik (2004) tested the consistency in coordination of the lip aperture (higher-level) and the lower-lip–jaw (lower level) synergy over repeated productions and found lower variability in the lip aperture synergy, suggesting its role as a functional synergy critical in determining speech acoustics. Based on theories of motor learning that predict progressive changes in spatio-temporal features of movement control and reduced movement variability (e.g., Kleim et al., 2003, 2004), and psycholinguistic research predicting improved phonemic encoding (e.g., Bishop, Barry, & Hardiman, 2012) with practice, we hypothesized that both groups will demonstrate an increase in stability of the lip aperture trajectory and a reduction in speech errors with practice. In addition, we hypothesized that participants in both groups will retain the encoded phonological information and the learned movement sequences in memory and therefore, show improvements in Session 2 compared to Session 1. Based on earlier studies (e.g., Ludlow et al., 2004; Namasivayam & van Lieshout, 2008; Smith et al., 2010; Smits-Bandstra et al., 2006), we hypothesized that the AWS will show limited benefits of practice and retention on speech errors and speech kinematics.

Performance in the nonword repetition task was assessed to answer the following research questions:

Behavioral

- a. Do AWS differ from controls in their responses to manipulations of nonword type (length, PC) and complexity?
- b. Do AWS differ from controls in the improvements in speech accuracy from Session 1 to 2?

Kinematic

- a. Do AWS differ from controls in their speech kinematic responses to manipulations of length, PC, and complexity?
- b. Do AWS differ from controls in speech kinematics (inter-articulator coordination, speech duration) with practice (within Session 1, first 5 vs. next 5 trials) and retention (first 5 trials, Session 1 vs. First 5 trials, Session 2; Next 5 trials Session 1 vs. first 5 trials, Session 2)?

Methods

Participants

Participants were 10 adults who stutter (9 males, 1 female, *Mean* age = 30.7 years, Age range = 17.7 – 65.1 years, *SD* = 15.1) and 10 age and sex matched typically fluent adults (*Mean* age = 30.8 years, Age range = 18.6 – 64.5 years, *SD* = 14.9). Each participant in the AWS group was matched within a one year age range to a participant in the control group. The groups were comparable in the mean number of years of post-secondary education (AWS, *Mean* = 5.1 years; *SD* = 2.8; Control, *Mean* = 4.4; *SD* = 3.0, $t(18) = -0.52$, $p = 0.6$). Of the participants included in the data set, one participant in each group was a 65 year old, the remaining participants were between 17.7 and 45 years. The decision to include the 65-year-olds in the data set is justifiable based on recent data demonstrating comparable kinematic performances between adults in this age group and younger adults (Sadagopan & Smith, In press). All participants, except one in the stuttering group, were monolingual native speakers of American English; the non-native speaker had native-like proficiency in English. Participants were recruited from a pre-existing database maintained at the Speech Fluency lab and through the Julia Davis Speech and Hearing clinic at the Department of Speech-Language-Hearing Sciences, University of Minnesota. The experimental protocol was approved by the Institutional Review Board, University of Minnesota.

Initial screening—Participant selection was based on responses to a screening form used to rule out positive history of language, hearing, and/or neurological deficits, and current usage of drugs likely to affect the outcome of the experiment (e.g. drugs for ADHD and anti-anxiety drugs). All participants passed a binaural hearing screening at 20 dB HL at .5, 1, 2, 4, and 8 KHz. Normal articulatory structures and movements in both groups were confirmed using the *Oral Speech Mechanism Screening Evaluation-Revised* (OSMSE-R; St.

Louis & Ruscello, 1987). All participants in the AWS group had a positive treatment history. Three of the AWS were enrolled in treatment at the time of participation.

Stuttering severity—Reading and speech samples collected from individuals in the AWS group were used to calculate stuttering severity using the Stuttering Severity Instrument – 3 (SSI – 3; Riley, 1994). A trained research assistant (RA I) coded the stuttered disfluencies (sound and syllable repetitions, monosyllabic word repetitions, prolongations, and blocks) in the samples. Accordingly, there were 6 individuals with very mild, 1 mild, 2 moderate, and 1 with severe SSI scores. For six of the participants, SSI severity ratings from the initial coding by RAI were compared to severity ratings obtained from a second RA (RAII) and resulted in 100% agreement between the two raters.

Standardized tests

A series of tests were administered to evaluate expressive vocabulary, lexical encoding, phonological encoding, semantic, syntactic encoding, and working memory abilities of participants at baseline in both groups. These skills were assessed due to their potential contributions to nonword repetition performance (e.g., Gupta, 2003). Expressive vocabulary was tested using the *Expressive Vocabulary test* (EVT; Williams, 1997). Lexical encoding and working memory were tested using the forward and backward digit span subtest of the *Wechsler's Intelligence test* (Wechsler, 1997). Semantic, syntactic encoding and working memory were tested using the Sentence Recall subtest of the *Clinical Evaluation of Language Fundamentals test - 3* (CELF - 3; Semel, Wiig, & Secord, 1995). Participants in both groups were also tested in the Nonword Repetition task (NRT; Dollaghan & Campbell, 1998) as a baseline measure of phonological encoding and working memory abilities. The NRT (1 – 4 syllable nonwords) consists primarily of tense vowels and consonants acquired early in development while excluding the late-eight consonants and consonant clusters (Dollaghan & Campbell, 1998).

Stimuli

For the purposes of this study, nonword type and complexity were varied to create the target nonwords. Nonword type was varied based on syllable length and PC (3-PC, 4, 6 syllable, 3-NPC). Each nonword type consisted of two nonwords classified as simple and complex based on nonword complexity. Complexity was operationally defined based on syllable-internal phonemic composition - early vs. middle, late consonants (Shriberg et al., 2010) and the number, nature of consonant clusters. The simple nonwords consisted of the early and middle 8 consonants and one or two bi-consonant clusters, while the complex nonwords consisted of the middle and late 8 consonants and three or more bi- and tri-consonant clusters (for a description of stimuli see, Table 1). Nonwords consisting of the middle and late 8 speech sounds as well as consonant clusters have been reported to elicit more speech errors and reduced inter-articulator coordination in both typically fluent children and adults (e.g., Sasisekaran, Smith, Sadagopan, & Weber-Fox, 2010; Sasisekaran & Weisberg, 2013). In addition, nonwords consisting of phonemic combinations less frequently or never encountered in a language (NPC) are likely to constraint the efficiency of retrieval, encoding, and production of such novel combinations (e.g., Bishop et al., 2012; Sasisekaran & Weisberg, 2013; Vitevich, Luce, Charles-Luce, & Kemmerer, 1997).

All nonwords started with /mæb/ and ended in a voiced bilabial (either /m/ or /b/). This strategy was used to select consistent start and end points for the extraction of oral movement trajectories on the basis of lower lip peak opening velocities for the kinematic analysis. All nonwords also carried an alternating strong – weak stress pattern.

Apparatus

Kinematic protocol—Participants were seated in front of an Optotrak Certus camera (Northern Digital, Waterloo, Ontario). This system allows tracking of movements in 3D with an accuracy of less than 0.1mm. Eight small (7 mm) infra-red light emitting diodes (IREDs) were attached to the participant's face to track articulatory movements of the upper lip, lower lip, and jaw. Four of the IREDs were mounted on a set of goggles that participants were required to wear during the experimental session. One IRED was placed in the center of the forehead. Together, these five IREDs were used to calculate the 3D head coordinate system (Smith, Johnson, McGillem, & Goffman, 2000), which allowed head movement artifact to be eliminated. The remaining three IREDs were placed one on the vermilion border of the upper lip, one on the center of the lower lip (this marker represents combined actions of the lower lip and jaw), and one on a splint taped to the jaw. The IREDs were fixed to the subject and facing the Optotrak camera such that the movements could be tracked. The IRED motions were sampled at a rate of 250 Hz. A wireless condenser microphone was placed approximately 8 cms away from the participant's mouth and a microphone receiver was used to record the acoustic speech signal. This acoustic signal was digitized on an A/D channel of the Optotrak system and was thus synchronized with the movement data. The acoustic signal was digitized at the rate of 16,000 Hz.

Procedure

The entire test session took about 2.5 hours and was divided into two Optotrak sessions (Sessions 1 and 2) separated by an hour of administering standardized tests. Practice effects were tested in Session 1. Retention was tested after an hour based on studies in the speech motor literature that have tested retention effects within this time frame (e.g., Behrman, et al., 2000; Smits-Bandstra et al., 2006). During both sessions, the nonwords, pre-recorded as spoken by a female native English speaker, were presented over loudspeakers. The principal investigator (PI) and a trained research assistant (RAI) were present throughout Sessions 1 and 2. While RAI presented the nonword stimuli from a laptop, the PI managed the Optotrak recording. Participants heard the target nonwords across several blocks and were required to repeat each nonword as accurately as possible. The nonwords repeated by the participants were recorded and analyzed offline to obtain percent speech accuracy across the different nonword types and complexity.

Prior to both sessions, participants were instructed that they would hear the eight target nonwords in a random order in several blocks and they were required to repeat the nonwords at comfortable loudness upon hearing each nonword. Following instructions and only for Session 1, participants were provided practice trials. During practice, participants were presented each nonword twice both in written format on a computer screen as well as aurally. Productions during practice were monitored and corrected for pronunciation and stress errors. The data indicated that a majority of the participants were able to achieve a

minimum of one correct production within the two practice blocks for the 3, 4, and 6 syllable nonwords. For the 3-NPC nonwords only, additional two blocks of practice trials were presented (if required) such that participants were able to produce the nonwords correctly at least once with assistance. Following practice, participants were instructed that they would hear the nonwords through the loudspeakers and were required to repeat the nonwords as accurately as possible.

For Session 1, a total of 19 to 22 blocks (excluding the practice blocks) were presented to all participants. For Session 2, a total of 10 blocks were presented. These many number of trials were chosen for practice and retention based on findings from earlier studies investigating such effects for nonwords varying in length and/or complexity in typically fluent adults and AWS (e.g., Bishop et al., 2012; Sasisekaran & Weisberg, 2013; Smith et al., 2010). Each block consisted of the eight test nonwords presented once in random order. In instances where additional blocks were required for a nonword type or complexity category, a few additional blocks were presented such that all participants were able to obtain the required number of correct productions of each nonword. Attempts were made to obtain the first set of 8 – 10 correct productions from Session 1 and the first set of 4 – 5 correct productions from Session 2 as these minimum number of trials were required for kinematic analysis. However, these correct productions were not necessarily consecutive correct productions for all participants and therefore, the individual trial numbers (Session 1: 1 – 22, Session 2, 1 – 10) included in the analysis for each participant varied. The 8 – 10 correct productions from Session 1 and the 4 – 5 correct productions from Session 2 were analyzed in three sets of 4 – 5 thus allowing a comparison of the first 4 – 5 (hereafter S1-F5) with the next 4 – 5 trials (hereafter S1-N5) from Session 1, and of S1-F5, S1-N5 with the first 4–5 trials from Session 2 (hereafter S2-F5). The number of trials chosen were based on earlier studies of movement variability that have used at least five early and five later trials to investigate practice effects in typically fluent and disfluent speakers (e.g., Sasisekaran, et al., 2010; Smith et al., 2010; Walsh et al., 2006).

During the experiment, the PI and RAI recorded participant errors and disfluencies online. Post-experiment, RAI listened independently to the recordings and provided offline scores of correct and error productions as well as disfluencies for the behavioral and kinematic analysis. Online coding of the productions by the PI and offline coding by RAI were later compared for inter-judge reliability. Only those productions that both the PI and RAI (offline and online) agreed upon as being correct and fluent were included in the kinematic analysis. In addition, a trained RA (RAIII) provided inter-judge reliability for the offline analysis done by RAI.

Data Analysis

Behavioral analysis—Nonwords produced with consonant or vowel errors were coded from the offline analysis. Errors included substitutions, omissions, additions, and distortions. Stress errors were not coded although observation revealed that a majority of participants were able to imitate the rhythm of the speaker. Disfluencies, including interjections, hesitations, sound or syllable repetitions, prolongations, and blocks, and self-corrects were included in the behavioral analysis as correct responses if participants did produce the

nonwords correctly. However, disfluent trials were not included in the kinematic analysis. In addition, productions with pauses (more than 1 sec) were also excluded from the kinematic analysis. Ratings from the offline analysis done by RAs I and III were compared. The two raters disagreed on a total of 0.07% ($N = 12$) trials in the AWS group and 0.07% trials ($N = 11$) in the control group, and therefore these trials were excluded from further analysis. Furthermore, six trials in the AWS group and three trials in the control group from Session 1 were excluded from behavioral and kinematic analyses due to ambiguity in classifying the productions as errors.

Kinematic analysis—A detailed description of this analysis can be obtained from Smith et al. (2000) and Smith and Zelaznik (2004). The MATLAB digital signal processing software was used to simultaneously load and analyze the acoustic and kinematic data from Sessions 1 and 2. The lip displacement signals (both upper and lower lip) were low pass filtered (cut-off 10 Hz) in both forward and backward directions using a Butterworth filter. Following this, the velocity signal was computed from the lower lip displacement signal using a three-point difference method. Following this, for Session 1, the lower lip displacement trajectories associated with the first 8 – 10 correct productions of each nonword at each nonword type and complexity were obtained by identifying the correct productions from the blocks and segmenting the corresponding displacement file for each correct production. An identical procedure was adapted for the first 4 – 5 correct productions from Session 2. The velocity signal was used to obtain the start and end point of the lower lip displacement trajectory of each correct production. The start point for each segment was the point of peak velocity of the lower lip opening gesture for /m/ (in the first syllable /mæb/) and the end point was the point of peak velocity of the lower lip opening gesture for /m/ or /b/ (in the last syllable). The peak velocity regions were easily located by the experimenter, and a MATLAB algorithm was used to select the peak opening velocity within a user-defined window. Each trim corresponding to a correct production selected in the above-described manner was then reassessed for fluency and accuracy of extraction by listening to the associated audio recording.

Following extraction of the lip displacement trajectories for a set of *S1-F5*, *S1-N5*, and *S2-F5* correct productions corresponding to each nonword, a multi-step analysis was performed using a custom MATLAB program. First, the lip aperture (LA) signal was obtained by a sample-by-sample subtraction of the upper lip from the lower lip superior - inferior displacement signal. This analysis was carried out to obtain the difference between the upper and lower lip IRED markers as a function of time. Second, the three sets of 4 – 5 LA trajectories for each nonword were amplitude normalized, which involved subtracting the *Mean* and dividing by the standard deviation. Third, using interpolation, the LA trajectories were time normalized to a fixed record length of 1000 points. Fourth, standard deviation values were calculated at 50 point intervals (2% intervals in relative time) of the normalized waveform. The cumulative sum of the 50 standard deviations was computed to obtain a cumulative spatial and temporal coordinative index called lip aperture variability index (LA VAR; Smith & Zelaznik, 2004). The LA VAR reflects the degree of spatial and temporal variability in coordinative patterns among the upper lip, lower lip, and jaw trajectories for each nonword production. A higher LA VAR index suggests a greater degree of trial to trial

variability in inter-articulatory coordination for nonword production, and vice versa. In addition to LA VAR, the speech duration in real time was computed as the total duration of the LA signal for each target production and was obtained by measuring the duration of each movement trajectory of a nonword from start to end point. Figure 1 illustrates the LA trajectories for the 6-syllable simple nonword, /mæbraibapotiba/ from Sessions 1 and 2 for a control participant (upper half) and an AWS (lower half).

Statistical Analysis

Behavioral data—We used a logistic generalized linear mixed model using the lme4 package in R (McCulloch, Searle & Neuhaus, 2008) with a binomial response, either correct or incorrect on each trial. Each subject was tested repeatedly and subject was a random effect. Group (AWS, Control), Nonword type (PC, 4, 6, 3-NPC), Complexity (simple vs. complex), and Session (1 vs. 2) were treated as factors and covariates were also fit. The covariates were average scores from EVT, digit span, sentence recall (%), and NRT (%). For the NRT task, percent correct productions averaged across the four nonword lengths was included as a covariate based on comparable performance between groups at each of these nonword lengths in this study and others (e.g., Smith et al., 2010). All interactions among the factors up to third order (four factor interactions had power that was too low to be meaningful), including main effects for the covariates and a random subject effect, were considered. Type II analyses of deviance chi-squared (χ^2) tests were used to determine significance of terms. As is standard (e.g., Fox & Weisberg, 2011, page 195), we tested the highest order interactions first and then the next level of interactions. To simplify the discussion of results we refit excluding the non-significant interaction (Group x Nonword type x Session) and discuss the results from the fitted model without this interaction. The null hypothesis tested was that a 3-way interaction of the factors is zero given all other terms, that is, there are no significant differences between groups based on nonword type and session. The alternate hypothesis was that the 3-way interaction is not equal to zero (i.e., the Group x Nonword type x Session interaction is significant), in which case any lower-level interactions (2-way interactions including Group x Nonword type, Group x Session, Nonword type x Session) may not be tested. Post-hoc comparisons were done using the lsmeans package in R, which uses Tukey's *highly significant difference* (Lenth, 2013). The aims of the mixed method analysis were: a) to study the effects of nonword type (3-PC, 4, 6, 3-NPC) and complexity (simple vs. complex) on the estimated probability of correct responses in AWS and control groups; b) to study group differences in the extent of changes in speech accuracy between sessions; and c) to investigate the effects of the covariates on nonword repetition performance.

Kinematic data—Due to an unequal number of participants contributing to the different nonwords, the factor Nonword type (3-PC, 4, 6, 3-NPC) was subdivided into Syllable length (3-PC, 4, 6) and PC (3-PC, 3-NPC) for further analysis. Table 2 shows the number of participants who achieved the criterion of 4 – 5 correct productions at each length, PC, and complexity for Sessions 1 and 2. Session 2 kinematic data from two AWS could not be recorded due to technical difficulties.

Movement variability: Length and complexity effects: The LA VAR scores were analyzed in two repeated measures ANOVA. The aims of these analyses were to study the effects of the 4 factors - Length, PC, Complexity, and Session - on movement variability in the AWS and control groups. For the first analysis, Group (AWS, Control) was the between-subject variable while Nonword length (3-PC, 4 syllables), Complexity (simple vs. complex), and Trail (S1-F5, S1-N5; S2-F5) were the within-subject variables. The 6-syllable nonwords were not included in this analysis as an unequal number of AWS and controls produced the 6-syllable *complex* nonword (see Table 2). In contrast, an approximately comparable number of participants from both groups were able to repeat the 6-syllable *simple* nonword. Therefore, a separate repeated measures ANOVA was run on the 6-syllable simple nonword with Group as the between- and Trial (S1-F5, S1-N5; S2-F5) as the within-subject factor. All post-hoc comparisons were done using Fisher's *least significant difference* (LSD). Finally, paired *t* tests were used to analyze the LA VAR scores between S1-F5 vs. S2-F5 trials from the control group only for the 6-syllable complex nonword. A similar comparison was not done in the AWS group as only a limited number of subjects were able to produce the required number of correct productions.

Phonotactic constraint: Due to the unequal number of participants from the AWS and control groups who were able to produce the required number of correct productions, repeated measures ANOVA could not be conducted to compare directly the effects of 3-PC vs. 3-NPC on movement variability. However, paired *t*-tests were run within groups to investigate the changes in LA VAR with practice (S1-F5 vs. S1-N5) for the simple and complex 3-NPC nonwords. Independent samples *t*-tests were also run to compare the groups' LA VAR scores only for the first 5 trials of the 3-NPC simple and complex nonwords as an equal number of participants from each group contributed to this data set (see Table 2).

Nonword duration: The duration data were analyzed separately in analyses similar to those described for the movement variability data.

Results

Standardized tests results

Table 3 presents the *Mean* and *SD* for expressive vocabulary (EVT), digit span (forward and backward), and sentence recall tests from both groups. Independent *t*-tests comparing the groups in EVT Standard scores revealed that the AWS scored significantly lower than the control group. The groups were comparable in all other tests. Repeated measures ANOVA of the NRT data did not show a significant Group effect, $F(1, 18) = 0.18, p = 0.67$. A significant main effect of Nonword length, $F(3, 54) = 31.8, p < 0.001$, *partial eta squared* (η_p^2) = 0.63, and subsequent post-hoc comparisons (*Fischer's LSD*) indicated a lower percent of correct phonemes at the 3- and 4-syllable lengths compared to the 1- (3 vs. 1 syllable, $p = 0.03$; 4 vs. 1 syllable, $p < 0.001$) and 2-syllable (3 vs. 2 syllable, $p = 0.01$; 4 vs. 2 syllable, $p < 0.001$) nonwords. The Group x Length interaction was not significant, $F(3, 54) = 1.8, p = 0.14$.

Behavioral Data

One subject in the AWS group received three additional blocks (22 total blocks) and one subject in the control group received one additional block (20 total blocks) in Session 1. The groups were comparable in the mean number of blocks presented, $t(18) = 0.63, p = 0.53$. The *Analysis of Deviance* results are presented in Table 4. The 3-way interaction of Group x Nonword type x Complexity was not significant, $\chi^2(3) = 6.7, p = 0.08$ (see Figure 2). Post-hoc analysis of the significant Group x Nonword type interaction showed that averaged across the simple and complex nonwords, the AWS and control groups were comparable in the estimated probability of correct responses for the 3-PC ($p = 0.62$) and 4-syllable ($p = 0.18$) nonwords. However, the AWS showed a significantly lower probability of correct responses for the 6-syllable ($p = 0.01$) and the 3-NPC nonwords ($p = 0.006$). Post-hoc analysis of the significant Group x Complexity interaction, $\chi^2(1) = 4.9, p = 0.026$, revealed that averaged across Nonword type the groups were comparable in the estimated probability of correct responses for the simple nonwords ($p = 0.15$) while the AWS group showed a significantly lower probability of correct responses for the complex nonwords ($p = 0.01$).

The significant Complexity x Nonword type x Session interaction showed that for the simple nonwords, across both groups a higher estimated probability of correct responses was seen for Session 2 compared to 1 for all nonwords ($p < 0.001$), except the 3-PC syllable simple nonword ($p = 0.12$). For the complex nonwords, a significantly higher estimated probability of correct response was seen for Session 2 compared to 1 for all categories ($p < 0.001$).

Covariate effects—As shown in Table 4, a significant effect of digit span on task performance was obtained. None of the other covariates, including expressive vocabulary, sentence recall, and performance in the NRT had a significant effect.

Kinematic Data

Movement variability, length and complexity effects—This repeated measures ANOVA was conducted with the 3-PC and 4-syllable simple and complex nonwords only. Mauchley's test of sphericity revealed p values of > 0.05 for the repeated measures. Therefore, sphericity assumed p values are reported. This analysis did not show a significant main effect of Group, $F(1, 11) = 0.8, p = 0.39$, or Complexity, $F(1, 11) = 1.7, p = 0.21$. Significant main effects of Nonword length, $F(1, 11) = 44.4, p < 0.001, \eta_p^2 = 0.80$, and Trial, $F(2, 22) = 31.3, p < 0.001, \eta_p^2 = 0.74$, were obtained. Across both groups, movement variability reduced significantly from: a) 4-syllable to 3-PC syllable nonwords ($p < 0.001$), and b) within (S1-F5 vs. S1-N5, $p < 0.001$), and between sessions (S1-F5 vs. S2-F5, $p < 0.001$). A significant Length x Complexity interaction, $F(1, 11) = 5.2, p = 0.04$, showed comparable LA VAR scores for the 3-syllable simple and complex nonwords ($p = 0.90$) while a significantly higher mean LA VAR score was observed for the 4-syllable complex nonword ($p = 0.01$). The Group x Complexity interaction was not significant, $F(1, 11) = 3.9, p = 0.07$. A significant Group x Length x Trial interaction was observed, $F(2, 22) = 22.8, p < 0.001, \eta_p^2 = 0.67$ (Figure 4). At the 3-syllable level, only the AWS showed a significant reduction in LA VAR within Session 1 ($p < 0.001$). Neither groups showed a significant reduction in LA VAR between sessions. The groups were not significantly different in LA VAR scores at the end of Session 2. At the 4-syllable level, the groups

showed higher and comparable LA VAR scores initially (S1-F5, $p = 0.57$). However, with practice the control group exhibited a significant reduction in movement variability within session ($p < 0.001$), and maintained this reduction between sessions (S1-N5 vs. S2-F5, $p = 0.15$). The AWS group did not show any reduction in movement variability (within, $p = 0.72$; between, $p = 0.38$). The 3-way interaction of Group x Complexity x Trail effect was not significant, $F(2, 22) = 2.63$, $p = 0.09$.

A repeated measures ANOVA was run on the LA VAR scores from the 6-syllable simple nonword. Mauchley's test of sphericity revealed a p value of > 0.05 . This analysis did not reveal a significant group effect, $F(1, 10) = 0.31$, $p = 0.58$. A main effect of Trial, $F(2, 20) = 1.04$, $p < 0.001$, $\eta_p^2 = 0.50$, showed a significant reduction in LA VAR from the first 5 to the next 5 trials within Session 1 ($p < 0.001$), and to the first 5 trials in Session 2 ($p = 0.001$), in both groups. A significant Group x Trial effect was not obtained, $F(2, 20) = 1.04$, $p = 0.37$.

Paired t-test was used to compare the LA VAR scores between sessions, from S1-F5 ($N = 8$) to S2-F5 trials ($N = 7$), for the complex 6-syllable nonword in the control group. Changes in LA VAR within session (i.e., S1-F5 vs. S1-N5) was not investigated as only four participants could achieve the required number of correct productions. A significance change in the mean LA VAR value was not observed in the control group for this nonword (S1-F5, $Mean = 27.9$, $SD = 4.6$; S2-F5, $Mean = 29.1$, $SD = 5.1$; $t(6) = 0.63$, $p = 0.54$). A similar analysis was not done in AWS as the required number of correct productions was not available.

Movement variability, Phonotactic constraint—Table 5 shows the *Mean* (and *SD*) LA VAR scores for the 3-PC and 3-NPC simple and complex nonwords from both groups. Although a direct comparison of the scores was not made, note the large difference in scores between the PC and NPC categories. Paired t-tests (*Bonferroni*-corrected for three comparisons, $p = 0.017$) of the LA VAR scores for the 3-NPC simple nonword in the control group showed only a descriptive reduction in LA VAR within session, $t(7) = 2.8$, $p = 0.02$, and significant difference between sessions (S1-F5 to S2-F5, $t(8) = 3.6$, $p = 0.008$). Practice effects in LA VAR were not evident for the complex 3-NPC nonword. Similar comparisons in the AWS group did not reveal any practice effects for the simple or the complex 3-NPC nonwords. Independent samples t-test comparison of the LA VAR scores from the first 5 trials only for the 3-NPC simple nonword, where the control group showed a descriptively higher mean LA VAR score, revealed a significant difference between the groups, $t(14) = -2.15$, $p = 0.04$. Independent samples t-test comparison of the LA VAR scores from the first 5 trials only for the 3-NPC complex nonword, revealed comparable group performances, $t(12) = 0.008$, $p = 0.9$.

Movement duration, Length and complexity effects—A repeated measures ANOVA was done on speech duration for the 3- and 4-syllable nonwords. Mauchley's test of sphericity revealed p values of < 0.05 for the Length effect. *Hyuen-Feldt* (*H-F*) p values are reported for the main and interaction effects involving this repeated measure. The repeated measures ANOVA did not show an overall significant Group effect, $F(1, 9) = 2.0$, $p = 0.18$. Significant main effects of Complexity, $F(1, 9) = 35.5$, $p < 0.001$, $\eta_p^2 = 0.79$,

Nonword length, $F(1, 9) = 89.9$, $(H - F) p < 0.001$, $\eta_p^2 = 0.90$, and Trial, $F(2, 18) = 5.7$, $p = 0.01$, $\eta_p^2 = 0.39$, were observed. Across both groups speech duration reduced from: a) complex to simple nonwords ($p < 0.001$), b) 4-syllable to 3-syllable nonwords ($p < .001$), and c) the first 5 trials in Session 1 to first 5 trials in Session 2 ($p = 0.02$). A significant Length x Complexity interaction was obtained, $F(1, 9) = 5.6$, $(H - F) p = 0.04$, $\eta_p^2 = 0.38$. All other interactions were not significant: Length x Trial, $F(2, 18) = 0.84$, $(H - F) p = 0.44$, $\eta_p^2 = 0.23$; Complexity x Trial, $F(2, 18) = 1.2$, $p = 0.32$; Group x Complexity, $F(1, 9) = 0.53$, $p = 0.48$; Group x Trial, $F(2, 18) = 2.7$, $p = 0.09$; Group x Length x Complexity, $F(1, 9) = 0.20$, $(H - F) p = 0.65$; Group x Length x Trial, $F(2, 18) = 0.72$, $(H - F) p = 0.49$; Group x Complexity x Trial, $F(2, 18) = 1.9$, $p = 0.17$; Length x Complexity x Trial, $F(2, 18) = 1.5$, $(H - F) p = 0.24$; Group x Length x Complexity x Trial, $F(2, 18) = 2.4$, $(H - F) p = 0.11$.

Repeated measures ANOVA of the speech duration for the simple 6-syllable nonword revealed a significant main effect of Trial within session, $F(2, 20) = 7.9$, $p = 0.002$, $\eta_p^2 = 0.44$. All other effects were not significant; Group (1, 10) = 0.90, $p = 0.36$; Group x Trial, $F(2, 20) = 1.5$, $p = 0.23$.

Movement duration, Phonotactic constraint—Table 5 shows the *Mean* (and *SD*) nonword duration for the 3-PC and 3-NPC simple and complex nonwords for both groups. Paired t-tests (*Bonferroni*-corrected for 3 comparisons, $p = 0.017$) revealed that the control group did not show any reduction in speech duration within Session 1, $t(6) = 2.1$, $p = 0.08$, or between Session 1 vs. 2, $t(7) = -0.73$, $p = 0.49$, for the 3-NPC simple nonword. A significant reduction in speech duration was observed between sessions (S1-F5 vs. S2-F5) for the complex 3-NPC nonword, $t(7) = 3.6$, $p = 0.01$. The AWS group did not show any significant changes in speech duration for the 3-NPC nonwords.

Discussion

The aim of the present study was to investigate skill acquisition in AWS using a nonword repetition task. AWS and control groups were compared in speech accuracy and speech kinematics (lip aperture variability and speech duration). The effects of nonword type and complexity on speech accuracy and speech kinematics within and between sessions were tested. AWS showed significantly more speech errors for the 6 syllable and 3-NPC nonwords and for the complex compared to the simple nonwords. AWS also showed a within session reduction in movement variability (practice effect) for the 3-syllable nonwords while the control group was performing at ceiling at this length. For the 4 syllable nonwords, AWS did not demonstrate practice or retention effects on movement coordination that was evident in the control group, although the groups were comparable in duration for all nonwords. The findings support the combined use of behavioral and kinematic data to interpret difficulties experienced by AWS in nonword repetition and highlight the relevance of systematic manipulations of task complexity in understanding the extent of the skill acquisition deficit in AWS.

Behavioral findings

In the present study, speech accuracy was compared for nonwords varying in nonword type (length, PC) and complexity within and between sessions. This analysis revealed that the AWS showed more speech errors than the control group for the 6 syllable and the 3-NPC nonwords and for the complex compared to the simple nonwords. In addition, lesser number of AWS were able to achieve the criterion of 4 – 5 correct productions needed for the kinematic analysis, particularly for the 4-, 6-, and 3-NPC nonwords. These findings suggest that AWS experienced difficulties in the accurate production of nonwords of increasing length, PC, and complexity, which are interpreted to indicate difficulties in phonemic encoding and/or speech motor processes. Furthermore, the groups were comparable in the improvements in speech accuracy with retention in Session 2 thereby suggesting that the ability to hold the encoded information in memory, at least for a shorter duration, may be comparable between the groups.

Significant group differences in speech accuracy have been reported in a few other studies of nonword repetition (e.g., Byrd et al., 2012; Ludlow et al. 1997), while the lack of such differences were reported in several others (Bauerly & De Nil, 2011; Smith et al., 2010; Smits-Bandstra et al., 2006). Earlier studies have not investigated speech accuracy in nonwords that do not follow English PC. The complex nonwords used in this study, particularly the complex 6-syllable nonword, are comparable in internal composition and syllabic structure to the nonwords used in Ludlow et al. (1997), and in nonword length to Byrd et al. (2012). This may account for the similarities in the finding of significant group differences. Furthermore, the lack of group differences in the 3- and 4-syllable nonwords concur with similar findings reported in Smith et al. (2010). The findings also suggest that varying nonword length, PC, and complexity, as is the case in this study, was successful in revealing subtle difficulties experienced by AWS in nonword repetition.

Covariate effects—The results from the analysis also indicated that across both groups none of the covariates, except digit span, influenced task performance. This finding corroborates other similar reports (Sasisekaran & Weisberg, 2013). At the outset the AWS group did score significantly lower in expressive vocabulary than the control group. However, the mixed model analysis did not reveal expressive vocabulary to be a significant contributor to task performance. Additionally, research supports a role for vocabulary size in nonword learning in children (e.g. Gathercole & Baddeley, 1989; Storkel, 2001; Storkel, 2004) while short-term memory rather than long-term sub-lexical representations is considered critical in adults (e.g. Gaskell & Dumay, 2003). Therefore, although the groups differ in expressive vocabulary, we speculate this is unlikely to have contributed to the observed group differences in speech errors.

Digit span contributed significantly to task performance. The length of the items in the digit span tasks are comparable to the length of the nonwords in this study and therefore suggest that the working memory processes used in both the tasks to hold, manipulate, and repeat may be similar. Furthermore, the groups did not differ in digit span scores (forward span, $p = 0.45$; backward span, $p = 0.58$) thus suggesting that lower working memory capacities are correlated with increased speech errors, particularly for the longer and more complex

nonwords, in AWS to a similar degree compared to the controls. Finally, the groups were comparable in the NRT test, which could be interpreted as indicating that they have comparable phonological encoding and working memory capacities. However, we interpret this finding to suggest that the NRT task is not sufficiently challenging to reveal the subtle behavioral differences in nonword repetition abilities that exist between the two groups as demonstrated in the mixed model analysis.

Kinematic findings

Length and complexity effects—Analysis of changes in movement variability within and between sessions for the nonwords showed significant group differences that were influenced by nonword length. For the 3-PC nonwords, the AWS showed a significant practice effect seen as a reduction in movement variability within Session 1. The control group, however, did not show any changes in movement variability for this nonword and in fact seemed to be operating at ceiling in terms of inter-articulatory coordination even at Session 1. With increasing length (4-syllable nonwords), however, the control group showed increasing stability in inter-articulatory coordination within session and maintained this improvement with retention in Session 2. The AWS failed to show significant changes in movement variability with practice or retention. Therefore, as hypothesized based on theories of motor learning, both groups showed a reduction in movement variability, particularly within Session 1, that was further influenced by nonword length. Such changes may be attributable to a reduction in neuromotor noise with practice and to adjustments in spatio-temporal coordinates resulting in progressively optimal movement trajectories with practice and learning (Kleim, 2003, 2004). However, such effects were seen in AWS even for nonword lengths where the controls seem to be performing at ceiling (e.g., 3-syllable nonwords), thereby suggesting that the speech motor system may not be functioning at its optimal capacity for learning new speech strings in AWS. Similar findings have been reported in several other studies, although with extended practice and time to retention (Bauerly & De Nil, 2011; Namasivayam & van Lieshout, 2008; Smits-Bandstra et al., 2006). We tested such effects with reduced practice and time to retention and found similar group differences at least for the shorter and simpler nonwords. No significant practice or retention effects on movement variability were observed for the 6-syllable nonwords in either of the groups, which suggested that the amount of practice or time to retention may have been insufficient to induce learning related changes in speech motor physiology for these nonwords and supported the need for further investigation of the effects of longer practice and time to retention on the learning of novel phonemic strings in AWS.

The present findings did not reveal significant differences in movement coordination between the AWS and control groups for simple *vs.* complex nonwords. However, the absence of such an effect in this study does not necessarily indicate the lack thereof. For instance, Smith et al. (2010) reported higher movement variability for a 4-syllable complex nonword compared to a phonemically simpler 4-syllable nonword. Considering that in this study the AWS and control participants did not contribute equally to the kinematic dataset, particularly for the longer (6-syllable) complex and 3-NPC nonwords, we speculate that a larger experiment with more participants may well produce significant effects (note that the Group x Complexity effect is at a *p* value of 0.07).

Phonotactic constraint—As hypothesized, the 3-NPC nonwords were characterized by higher movement variability compared to the 3-PC nonwords in both groups. For the 3-NPC simple nonword, where the groups were comparable in accuracy, the control group showed significantly higher movement variability than the AWS for the first 5 trials. Although preliminary in nature, this finding is interpreted to suggest that in producing the simple NPC nonword the control group may have shifted to the use of a more flexible motor strategy, necessitated by the unfamiliar phonemic combinations. Similar findings of higher movement variability in control participants have also been reported at faster speech rates (Namasivayam & van Lieshout, 2008). With practice, the control group showed a significant reduction in movement variability and retained this gain in the second session, at least for the 3-NPC simple nonword. Neither did the AWS show the use of such a strategy nor did they show changes in movement variability with practice or retention. However, the argument for the use of a flexible movement strategy by the control group to produce the NPC nonwords is weakened by the fact that this strategy was not evident for the complex NPC nonword. Both groups showed comparable LA VAR scores for the first 5 trials of this nonword. Nonetheless, while eight of the ten participants in the control group were able to produce the minimum 4 – 5 correct productions necessary for kinematic analysis in Session 2 for this nonword, only three AWS were able to do so, which suggested continued difficulties in the latter group with this nonword.

Potential Underlying Sources of Behavioral vs. Kinematic Group Differences

Present findings support the combined use of behavioral and kinematic data to interpret the nature and extent of difficulties in nonword repetition in AWS. Based on the notion of the syllable as the minimal unit of motor control (e.g., Davis & MacNeilage, 2004), a higher percent of speech errors for the longer nonwords is, arguably, attributable to speech motor processes. A similar effect of complexity on task performance, particularly the 6-syllable and 3-NPC nonwords, is attributable to either phonemic encoding or speech motor processes. Such nonwords require activation, encoding, and execution of sound combinations that have previously not been (or have been less frequently) encountered and not stored in long-term memory. If speech motor processes were solely driving both behavioral and kinematic measures, group differences in both these measures should mirror each other at all levels of length, PC, and complexity. However, this was not the case and therefore, we speculate that the reduced speech accuracy for complex nonwords in AWS is attributable to not just speech motor processes, such as, breakdown in articulation, but also to difficulties in phonemic encoding. Yet another finding of interest was that although, unlike the controls, the AWS failed to show any effects of retention on movement variability in Session 2, the groups were comparable in the reduction in speech errors in Session 2 with retention. This suggests potential differences in the retention of motor memory for movement sequences *vs.* memory of phonological spell-out in this group, although this finding warrants further follow-up.

Conclusions and Limitations

Present findings support the presence of a speech skill acquisition deficit in AWS. Group differences in both speech accuracy and speech kinematics were evident with nonword

repetition. Furthermore, manipulating nonword length, PC, and complexity resulted in a higher error rate in AWS that was not reported in all of the earlier studies. Reduction in speech errors across sessions (as a result of retention) was comparable in the two groups. The movement variability data demonstrated that AWS showed limited practice and retention effects on inter-articulatory coordination even for shorter and simpler nonwords where the groups were comparable in speech errors.

Some cautions are raised in interpreting the present findings. First, fewer AWS than control participants contributed to some of the nonword categories. Second, only five trials were used to assess speech motor practice and learning effects. We justify this based on other studies that have used a similar number of trials and reported such effects with nonwords varying in length, PC, and complexity in both typically fluent individuals and in AWS (Sasisekaran et al., 2010; Smith et al., 2010; Walsh et al., 2006). Furthermore, using the same nonwords as those in this study, Sasisekaran and Weisberg (2013) reported similar findings in speech accuracy and movement variability in a group of monolingual and bilingual speakers. Finally, the present study investigated short-term skill learning in AWS using a limited set of nonwords, and the groups could not be compared in speech kinematics at all levels of nonword length and complexity. This necessitates further in-depth investigations of speech skill acquisition with extended practice and longer time to retention in AWS with a larger set of nonwords.

Acknowledgments

This study was funded by an NIH grant (NIDCD R03 DC010047). We express gratitude to our participants. We also acknowledge Brittany Graves, Jillian Stein, and Samantha Weaver for assistance with data collection and analysis, Dr. Edward Carney for technical assistance, and Dr. Anne Smith for feedback on the final version of this manuscript.

References

- Adam J. Historical Review and Appraisal of Research on the Learning, Retention, and Transfer of Human Motor Skills. *Psychological Bulletin*. 1987; 103(1):41–74.
- Adams MR, Hayden P. The ability of stutterers and nonstutterers to initiate and terminate phonation during production of an isolated vowel. *Journal of Speech and Hearing Research*. 1976; 19:290–296. [PubMed: 979203]
- Bauerly KR, De Nil LF. Speech sequence skill learning in adults who stutter. *Journal of Fluency Disorders*. 2011; 36:349–360. [PubMed: 22133413]
- Behrman AL, Cauraugh JH, Light KE. Practice as an intervention to improve speeded motor performance and motor learning in Parkinson's disease. *Journal of the Neurological Sciences*. 2000; 174:127–136. [PubMed: 10727698]
- Bishop DVM, Barry JG, Hardiman MJ. Delayed Retention of New Word-Forms Is Better in Children than Adults Regardless of Language Ability: A Factorial Two-Way Study. *PLoS ONE [Electronic Resource]*. 2012; 7(5):e37326.
- Byrd CT, Vallely M, Anderson JD, Sussman H. Nonword repetition and phoneme elision in adults who do and do not stutter. *Journal of Fluency Disorders*. 2012 Advance online publication. 10.1016/j.jfludis.2012.03.003
- Cohen J. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*. 1960; 20:37–46.
- Cooper MH, Allen GD. Timing control and accuracy in normal speakers and stutterers. *Journal of Speech and Hearing Research*. 1977; 20:55–71. [PubMed: 846204]

- Cross DE, Luper HL. Voice reaction time of stuttering and nonstuttering children and adults. *Journal of Fluency Disorders*. 1979; 4:59–77.
- Davis BL, MacNeilage PF. The frame/content theory of speech evolution: from lip smacks to Syllables. *Primates*. 2004; 6:305–328.
- Dollaghan C, Campbell TF. Nonword repetition and child language impairment. *Journal of Speech, Language, and Hearing Research*. 1998; 41:1136–1146.
- Fox, J.; Weisberg, S. *Companion to Applied Regression*. Sage; Thousand Oaks, CA: 2011.
- Gaskell GM, Dumay N. Lexical competition and the acquisition of novel words. *Cognition*. 2001; 89:105–132. [PubMed: 12915296]
- Gathercole SE. Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics*. 2006; 27:513–543.
- Gathercole SE, Baddeley AD. Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. *Journal of Memory and Language*. 1989; 28:200–213.
- Green JR, Moore CA, Reilly KJ. The sequential development of jaw and lip control for speech. *Journal of Speech Language and Hearing Research*. 2002; 45(1):66–79.
- Gupta P. Examining the relationship between word learning, nonword repetition, and immediate serial recall in adults. *Quarterly Journal of Experimental Psychology Section A- Human Experimental Psychology*. 2003; 56(7):1213–1236.
- Harris CM, Wolpert DM. Signal-dependent noise determines motor planning. *Nature*. 1998; 394:780–784. [PubMed: 9723616]
- 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.
- Kalvaram, KTh. Neurobiology of speaking and stuttering. In: Bosshardt, H-G.; Yaruss, JS.; Peters, HFM., editors. *Fluency disorders: Theory, research, treatment, and self-help; Proceedings of the 3rd World Congress of Fluency Disorders*; Nijmegen, The Netherlands: Nijmegen University Press; 2001. p. 59-65.
- Kleim JA, Barbay S, Cooper NR, Hogg TM, Reidel CN, Remple MS, et al. Motor learning-dependent synaptogenesis is localized to functionally reorganized motor cortex. *Neurobiology of Learning and Memory*. 2002; 77(1):63–77. [PubMed: 11749086]
- Kleim JA, Bruneau R, Calder K, Pocock D, VandenBerg PM, MacDonald E, et al. Functional organization of adult motor cortex is dependent upon continued protein synthesis. *Neuron*. 2003; 40(1):167–176. [PubMed: 14527441]
- Lenth, RV. R package (Version 1.06–05). 2003. Retrieved from <http://CRAN.R-project.org/package=lsmmeans>
- Ludlow, LC.; Siren, K.; Zikria, M. Speech production learning in adults with chronic developmental stuttering. In: Peters, HMF.; Hulstijn, W.; Starkweather, CW., editors. *Speech motor control and stuttering*. Amsterdam: Elsevier/Excerpta Medica; 1997.
- Luft AR, Buitrago MM. Stages of Motor Skill Learning. *Molecular Neurobiology*. 2005; 32:205–216. [PubMed: 16385137]
- McCulloch, CE.; Searle, SR.; Neuhaus, JM. *Generalized, Linear, and Mixed Models*. 2. Hoboken NJ: Wiley; 2008.
- Max L, Baldwin CJ. The role of motor learning in stuttering adaptation: Repeated versus novel utterances in a practice-retention paradigm. *Journal of Fluency Disorders*. 2010; 33(1):33–43. [PubMed: 20412981]
- Max L, Guenther FH, Gracco VL, Gosh SS, Wallace ME. Unstable or Insufficiently Activated Internal Models and Feedback-Biased Motor Control as Sources of Dysfluency: A Theoretical Model of Stuttering. *Contemporary Issues in Communication Sciences and Disorders*. 2004; 31:105–122.
- Monfils MH, Plautz EJ, Kleim JA. In search of the motor engram: Motor map plasticity as a mechanism for encoding motor experience. *Neuroscientist*. 2005; 11(5):471–483. [PubMed: 16151047]
- Nam, H. Syllable-level intergestural timing model: Split-gesture dynamics focusing on positional asymmetry and moraic structure. In: Cole, J.; Hualde, JI., editors. *Laboratory Phonology*. Berlin: Mouton de Gruyter; 2007. p. 483-506.

- Namasivayam AK, Van Lieshout PHHM. Investigating speech motor practice and learning in people who stutter. *Journal of Fluency Disorders*. 2008; 33:32–51. [PubMed: 18280868]
- Neilson, MD.; Neilson, PD. Adaptive model theory of speech motor control and stuttering. In: Peters, HFM.; Hulstijn, W.; Starkweather, CW., editors. *Speech motor control and stuttering*. Amsterdam, The Netherlands: Elsevier; 1991. p. 149–156.
- Newell KM, Liu YT, Mayer-Kress G. A dynamical systems interpretation of epigenetic landscapes for infant motor development. *Infant Behavior & Development*. 2003; 26(4):449–472.
- 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]
- Riley, G. A stuttering severity instrument for children and adults. 3. SSI. Austin, TX: Pro-Ed; 1994.
- Sadagopan N, Smith A. Age-Differences in Speech Motor Performance on a Novel Speech Task. *Journal of Speech, Language, and Hearing Research*. In press. 10.1044/1092-4388(2013/12-0293
- Sasisekaran J, Weisberg S. The effects of cognitive: linguistic variables and language experience on behavioral and kinematic performances in nonword learning. *Journal of Psycholinguistic Research*. 2013; 42(2):175–90. [PubMed: 22476630]
- Sasisekaran J, Smith A, Sadagopan N, Weber-Fox C. Nonword Repetition in Children and Adults: Effects on movement coordination. *Developmental Science*. 2010; 13(3):521–532. [PubMed: 20443972]
- Semel, S.; Wiig, EH.; Secord, WA. *Clinical Evaluation of Language Fundamentals - 3*. San Antonio, TX: Harcourt Brace & Company; 1996.
- Shriberg LD, Austin D, Lewis BA, McSweeney JL, Wilson DL. The Speech Disorders Classification System (SDCS): Extensions and lifespan reference data. *Journal of Speech, Language, and Hearing Research*. 1997; 40:723–740.
- Smith, A.; Kelly, E. Stuttering: A dynamic, multifactorial model. In: Curlee, RF.; Siegel, GM., editors. *Nature and treatment of stuttering: New directions*. 2. Needham Heights, MA: Allyn & Bacon; 1997. p. 204–217.
- Smith A, Zelaznik HN. Development of functional synergies for speech motor coordination in childhood and adolescence. *Developmental Psychobiology*. 2004; 45(1):22–33. [PubMed: 15229873]
- Smith A, Johnson M, McGillem C, Goffman L. On the assessment of stability and patterning of speech movements. *Journal of Speech Language and Hearing Research*. 2000; 43(1):277–286.
- Smith A, Sadagopan N, Walsh B, Weber-Fox C. Increasing phonological complexity reveals heightened instability in inter-articulatory coordination in adults who stutter. *Journal of Fluency Disorders*. 2010; 35(1):1–18. [PubMed: 20412979]
- Smits-Bandstra S, De Nil LF, Saint-Cyr JA. Speech and nonspeech sequence skill learning in adults who stutter. *Journal of Fluency Disorders*. 2006; 31:116–136. [PubMed: 16735062]
- StLouis, KO.; Ruscello, DM. *Oral Speech Mechanism Screening Evaluation -Revised*. Austin, TX: Pro-Ed; 1987.
- Storkel HL. Learning new words: Phonotactic probability in language development. *Journal of Speech, Language, and Hearing Research*. 2001; 44(6):1321–1337.
- Storkel HL. The emerging lexicon of children with phonological delays: Phonotactic constraints and probability in acquisition. *Journal of Speech, Language, and Hearing Research*. 2004; 47:1194–1212.
- Takahashi CD, Nemet D, Rose-Gottron CM, Larson JK, Cooper DM, Reinkensmeyer DJ. Neuromotor Noise Limits Motor Performance, But Not Motor Adaptation, in Children. *Journal of Neurophysiology*. 2002; 90:703–711. [PubMed: 12904490]
- Tilsen S, Goldstein L. Articulatory gestures are individually selected in production. *Journal of Phonetics*. 2012; 40:764–779. [PubMed: 23002316]
- Vangalen GP, Portier SJ, Smitsengelsman BCM, Schomaker LRB. Neuromotor noise and poor handwriting in children. *Acta Psychologica*. 1993; 82(1–3):161–178. [PubMed: 8475764]
- van Lieshout, PMMH.; Hulstijn, W.; Peters, HFM. Searching for the weak link in the speech production chain of people who stutter: A motor skill approach. In: Maassen, B.; Kent, R.; Peters,

- H.; van Lieshout, P.; Hulstijn, W., editors. Speech motor control in normal and disordered speech. Oxford: Oxford University Press; 2004. p. 313-356.
- Vitevich M, Luce P, Charles-Luce J, Kemmerer D. Phonotactics and syllable stress: Implications for the processing of spoken nonsense words. *Language and Speech*. 1996; 40:47–62.
- Walsh B, Smith A, Weber-Fox C. Short-term plasticity in children's speech motor systems. *Developmental Psychobiology*. 2006; 48(8):660–674. [PubMed: 17111401]
- Wechsler Memory Scale. 3. Published by The Psychological Corporation; 1997.
- Williams, KT. EVT-2: Expressive Vocabulary Test. 2. Bloomington, MN: Pearson; 1997.
- Willingham DB. A neuropsychological theory of motor skill learning. *Psychological Review*. 1998; 105:558–84. [PubMed: 9697430]
- Yan JH, Thomas JR, Stelmach GE, Thomas KT. Developmental features of rapid aiming arm movements across the lifespan. *Journal of Motor Behavior*. 2000; 32:121–140. [PubMed: 11005944]

Figure 1.

Lip aperture trajectories for /*mæbtaibapotiba*/ from Sessions 1 and 2 for a control (upper half) and an AWS participant (lower half). In the left-most panel, amplitude and time normalized trajectories for the first 5 trials are plotted, trajectories for the next 5 trials from Session 1 are plotted in the middle panel, and trajectories for the first 5 trials from Session 2 are plotted in the right-most panel.

Figure 2.

Estimated probability of correct productions (and 95% CI) by nonword type (length, PC) and complexity averaged across sessions for AWS and control groups.

Figure 3. Estimated probability of correct productions (and 95% CI) by nonword type (length, PC), complexity, and session averaged across AWS and control groups.

Figure 4.

Mean LA VAR scores (and *SD*) for the 3-PC, 4-, and 6-syllable (simple only) nonwords within (S1-F5, S1-N5) and between (S2-L5) sessions for AWS and control groups.

Table 1

Nonword stimuli by length, complexity, phonotactic constraint, and phonemic composition.

Nonword type				
Nonword	Length (in syll.)	PC/NPC	Complexity	Phonemic composition
/mæbbɪtiem/		PC	Simple	Early 8 consonants
/mæbbeiztʃab/		PC	Complex	Early, middle, late 8 consonants
/mæbspoukwiflɛɪb/	4	PC	Simple	Middle, late 8 consonants, 1 consonant cluster
/mæbskɪsploɪstrub/		PC	Complex	Middle, late 8 consonants, 3 consonant clusters
/mæbtɔɪbapotɪbɔ/	6	PC	Simple	Early, middle 8 consonants
/mæbgrɔɪroplɔkɪrɪbɔ/		PC	Complex	Early, middle, late 8 consonants, 4 consonant clusters
/mæbθwaɪpfkɪrɔb/	3	NPC	Simple	Early, middle, late 8 consonants, 2 consonant clusters
/mæbfʊdʒtʃloɪb/		NPC	Complex	Early, middle, late 8 consonants, 2 consonant clusters

Table 2
Number of participants contributing to LA VAR data for each nonword type by complexity for AWS and control groups.

Group	Total Number of subjects	3-PC		4 syllable		6 syllable		3-NPC	
		Trial	simple	complex	simple	complex	simple	complex	simple
AWS	10	S1-F5	10	10	9	9	7	4	7
	10	S1-N5	10	10	8	8	6	1	6
	8	S2-F5	7	7	6	5	5	2	4
Control	10	S1-F5	10	10	10	10	10	8	9
	10	S1-N5	10	10	10	9	8	4	7
	10	S2-F5	10	10	10	10	8	7	9

Table 3

Mean (and SD), *t*, *df*, and *p* values for EVT (Standard score), forward and backward digit span (raw score), sentence recall (%) for AWS and control groups.

Measure	AWS	Control	<i>t</i> value	<i>df</i>	<i>p</i> value
EVT	111.1	120.6			
	<i>SD</i> 6.5	9.4	-2.6	18	0.01*
Forward Digit span (Total = 16)					
	11.8	11.0			
	<i>SD</i> 1.7	2.3	0.8	18	0.45
Backward Digit span (Total = 14)					
	7.0	7.6			
	<i>SD</i> 2.9	1.6	-0.6	18	0.58
Sentence Recall					
	88.7	90.1			
	<i>SD</i> 5.6	9.4	-0.4	18	0.72

Table 4

Analysis of Deviance Table (Type II Wald chi-square tests, χ^2 , based on change in -2 likelihood) for the Logistic mixed methods analysis of the speech accuracy data.

FACTORS	χ^2	df	$p (> \chi^2)$
Complexity	126.5	1	<2e-16***
Nonword type	548.5	3	<2e-16***
Session	112.2	1	<2e-16***
Group	5.76	1	0.016*
Complexity x Nonword type	32.2	3	4.8e-07***
Complexity x Session	3.5	1	0.06
Nonword type x Session	4.1	3	0.24
Group x Complexity	4.9	1	0.026*
Group x Nonword type	8.9	3	0.03*
Complexity x Nonword type x Session	548.5	3	<2e-16***
Group x Complexity x Nonword type	6.7	3	0.081
COVARIATES			
Expressive Vocabulary	0.2	1	0.67
Sentence Recall	0.3	1	0.563
NRT	1.5	1	0.216
Digit Span	18.6	1	1.6e-05***

P values:

<.0001

**
<.01

*
<.05

<.1

Note. "Nonword type" includes nonword length and phonotactic constraint variables.

Table 5

Mean (and SD)

LA VAR (in millimeters) and speech duration (in s) for the simple and complex 3-PC and 3-NPC nonwords by group.

LA VAR	Simple 3-NPC		Complex 3-NPC	
	Mean	SD	Mean	SD
Control	First 5	31.8	5.3	24.9
	Next 5	23.3	6.4	20.4
	Later 5	21.4	7.9	24.3
AWS	First 5	24.1	8.8	24.9
	Next 5	22.3	10.8	22.8
	Later 5	24.1	8.7	22.6
Simple 3-PC				
Control	First 5	14.6	3.6	16.9
	Next 5	14.1	2.1	15.5
	Later 5	13.9	2.6	12.8
AWS	First 5	19.6	8.3	16.3
	Next 5	14.2	5.4	13.9
	Later 5	14.1	4.7	13.8
Complex 3-PC				
Duration	Simple 3-NPC		Complex 3-NPC	
	Mean	SD	Mean	SD
Control	First 5	1.8	0.3	2.2
	Next 5	1.7	0.5	2.0
	Later 5	1.8	0.5	1.9
AWS	First 5	1.7	0.4	2.2
	Next 5	1.9	0.8	2.2
	Later 5	1.6	0.1	1.9
Simple 3-PC				
Control	First 5	0.9	0.2	1.2
	Next 5	0.9	0.2	1.2
	Later 5	0.9	0.2	1.2

Duration	Simple 3-NPC		Complex 3-NPC	
	Mean	SD	Mean	SD
AWS	Next 5	0.9	0.1	1.1
				0.2
	Later 5	0.9	0.2	1.1
				0.2
	First 5	1.1	0.2	1.3
				0.3
	Next 5	1.0	0.1	1.2
				0.1
	Later 5	1.0	0.1	1.2
				0.2