Research Article

Do People With Apraxia of Speech and Aphasia Improve or Worsen Across Repeated Sequential Word Trials?

Katarina L. Haley,[a] Adam Jacks,[a] Jessica D. Richardson,[b] Tyson G. Harmon,[c] Elizabeth H. Lacey,[d] and Peter Turkeltaub[d]

[a] Division of Speech and Hearing Sciences, Department of Allied Health Sciences, The University of North Carolina at Chapel Hill [b] Department of Speech and Hearing Sciences, The University of New Mexico, Albuquerque [c] Department of Communication Disorders, Brigham Young University, Provo, UT [d] Department of Neurology, Georgetown University Medical Center, and MedStar National Rehabilitation Hospital, Washington, DC

ABSTRACT

Purpose: During motor speech examinations for suspected apraxia of speech (AOS), clients are routinely asked to repeat words several times sequentially. The purpose of this study was to understand the task in terms of the relationship among consecutive attempts. We asked to what extent phonemic accuracy changes across trials and whether the change is predicted by AOS diagnosis and sound production severity.

Method: One hundred thirty-three participants were assigned to four diagnostic groups based on quantitative metrics (aphasia plus AOS, aphasia-only, and aphasia with two borderline speech profiles). Each participant produced four multisyllabic words 5 times consecutively. These productions were audio-recorded and transcribed phonetically and then summarized as the proportion of target phonemes that was produced accurately. Nonparametric statistics were used to analyze percent change in accuracy from the first to the last production based on diagnostic group and a broad measure of speech sound accuracy.

Results: Accuracy on the repeated words deteriorated across trials for all groups, showing reduced accuracy from the first to the last repetition for 62% of participants. Although diagnostic groups differed on the broad measure of speech sound accuracy, severity classification based on this measure did not determine degree of deterioration on the repeated words task.

Discussion: Responding to a request to say multisyllabic words 5 times sequentially is challenging for people with aphasia with and without AOS, and as such, performance is prone to errors even with mild impairment. For most, the task does not encourage self-correction. Instead, it promotes errors, regardless of diagnosis, and is, therefore, useful for screening purposes.

Acquired apraxia of speech (AOS) is defined conceptually as a motor programming or planning disorder and behaviorally as a syndrome with a characteristic profile of abnormal speech sound and prosody production (Jacks & Haley, 2021; McNeil et al., 2009). The diagnosis is typically accompanied by aphasia, but occasionally, it presents with no or minimal language impairment. There

is lingering uncertainty about exactly how the behavioral speech profile of AOS differs from that of aphasia with phonemic paraphasia, which is conceptualized as a linguistic/phonologic impairment. This study is part of a larger project aiming to delineate behavioral diagnostic boundaries where they exist and identify speech tasks that are informative for severity estimation and treatment planning.

The study focused on a traditional task within typical motor speech examinations for AOS: consecutive repetition of multisyllabic words. The original purpose of the

task appears to have been to evaluate stability across repeated attempts with the expectation that people with AOS would demonstrate limited consistency. In their seminal book, Wertz and colleagues explained: "Repeated trials on the same word may show inconsistent articulatory errors and a combination of correct and incorrect productions among the five attempts" (Wertz et al., 1984, p. 103). In more recent reviews, different authors argued instead that production consistency is relatively strong in AOS, particularly in comparison to aphasia with phonemic paraphasia (McNeil et al., 2009; Wambaugh et al., 2006). A debate ensued about whether the original or revised consistency criteria should be used when diagnosing AOS. In the following, we first summarize what is known about production consistency on the task and then shift our focus to changes in proportional accuracy across sequential repetitions that are either self-initiated or requested.

Repeated Production to Reveal Inconsistency

On face value, a speaker who is asked to say difficult words multiple times should produce incorrectly some proportion of speech sounds, and any inconsistency among these errors would be evident to an attentive listener, assuming the errors themselves are perceptible. Indeed, diagnosticians have found the task useful in eliciting errors and have verified that word productions often do vary from attempt to attempt. The original clinical observations that production difficulties in AOS are inconsistent have stood the test of time for clinical diagnosticians (Molloy & Jagoe, 2019).

The reason that confusion emerged about production consistency in AOS is not found in clinical gestalt observation, but rather in how behavior has been measured, compared, and interpreted in research studies—all of which have been conducted with small sample sizes (ranging from four to 20 participants with AOS plus—in some cases—an approximately equally sized control group of people with aphasia and no AOS). When measuring error consistency, it is important to remember that it—like most behaviors—can be either consistent or inconsistent, depending on one's definition (Shuster & Wambaugh, 2008). On the one hand, error frequency and error patterns tend to be consistent in AOS (Mauszycki et al., 2010), allowing for reliable documentation of severity and a rationale for sound-based treatments. On the other hand, uncertainty remains regarding exactly how a challenging word will be produced (Haley & Martin, 2011), and variability increases with the unit of analysis as the field of potential variants increases from sounds to syllables to words (Haley et al., 2018).

Importantly, inconsistency of word retrieval and deficits in other linguistic behaviors and domains—including phonology—is also characteristic of aphasia (Goodglass, 1993). This means that differences in sound production consistency vis-a-vis AOS would, at a minimum, need to be expressed quantitatively and ideally relative to normative data. Because most people with stroke-induced AOS also have aphasia and because both linguistic and motor processes shape the speech output, it is an unavoidable reality that the relative contributions of linguistic and motor networks cannot be parsed during behavioral assessment. Consequently, one must ask the question whether the presence of AOS would alter production consistency beyond a baseline inconsistency of phonological origin—a question that would require a large and representative participant sample. To complicate matters further, it has been difficult to integrate results cross research groups for virtually all clinical research on AOS because diagnostic criteria have varied and diagnostic validity remains unverifiable due to reliance on clinical impression (Haley et al., 2012; Wambaugh et al., 2006).

To circumvent the complexity of interacting impairments at linguistic and motor levels, some researchers restricted their sampling to stroke survivors who have AOS and no more than minimal aphasia. In addition to severely limiting the feasible sample size, an unfortunate trade-off was that generalization of results became restricted to a presentation that is highly unusual. McNeil and colleagues compared sound error consistency for four people with relatively isolated AOS to four people with aphasia and phonemic paraphasia (McNeil et al., 1995). The analysis was based on three consecutive repetitions of 10 multisyllabic target words. On average, those with relatively isolated AOS produced more consistent errors than those with phonemic paraphasia. Based on the argument that cases of isolated AOS are diagnostically more informative than studies of AOS with coexisting aphasia, this study prompted a recommendation to alter existing diagnostic criteria from stating that sound errors in AOS are inconsistent to stating instead that AOS is characterized by sound errors that are relatively "consistent in terms of location and invariable in terms of type" (McNeil et al., 2009; Wambaugh et al., 2006).

Researchers continued to study sound production consistency in more diverse samples of left hemisphere stroke survivors with aphasia, and results challenged the altered diagnostic criterion (Bislick et al., 2017; Haley et al., 2013; Scholl et al., 2017; Staiger & Ziegler, 2008). In fact, where a difference existed, it suggested *lower*, rather than higher, consistency in AOS compared to aphasia and there were indications that consistency magnitude could be explained by impairment severity. In our group's

most recent study, speech and error consistency were examined in 137 participants with left hemisphere lesions due to stroke or trauma (Haley, Cunningham, Jacks, et al., 2021). Instead of diagnostic impression, objective measures were used to form four comparison groups. One group exhibited both slow multisyllabic word production and high sound distortion frequency and was considered to meet core diagnostic criteria for AOS; one group showed fast/normal multisyllabic word production and low sound distortion frequency and was considered to meet diagnostic criteria for aphasia only; two "borderline" groups had values that were intermediate between these profiles. Study results showed that all four groups produced multisyllabic words with varying degrees of inconsistency when asked to say them 5 times consecutively and that consistency measures, like previous smaller studies, were either similar or lower in the AOS group compared to the aphasia-only group. High speech sound error frequency predicted low error consistency at the word level (consistency of error type and production, due to more incorrect word variants) and high consistency in terms of what target segments were produced incorrectly (consistency of error location because more errors translate to more incorrect sound locations, regardless of the nature of the errors). The evidence showed strongly that relative consistencies of sound error location and type are not valid diagnostic criteria for differentiating between AOS and aphasia with phonemic paraphasia.

Having demonstrated that consistency metrics are of limited diagnostic use, the question remains whether the challenge of repeating words multiple times might help differential diagnosis in other ways. One previously considered possibility is that the task provides information about a speaker's ability to self-correct and that this ability could potentially inform diagnosis.

Sound-Level Self-Corrections in AOS and Aphasia

Self-corrections are common in AOS (Bailey et al., 2017; Harmon et al., 2019). Sometimes, they are heard as successive attempts referred to as "groping," a term that also encompasses inaudible articulatory posturing (Darley, 1968). Self-corrections are referenced in traditional accounts, stating that AOS involves "effortful, trial and error, groping articulatory movements and attempts at self-correction" (Wertz et al., 1984, p. 81). Darley and colleagues observed: "The apraxic patient effortfully gropes to find the correct articulatory postures and sequences of them. He often behaves as though uncertain of where his tongue is or how to move it in a given direction or to a given position" (Darley et al., 1975, p. 263). These

speaker-initiated successive attempts are referred to as false starts, revisions, re-approaches, or struggles (Johns & Darley, 1970; Liss, 1998; Odell et al., 1991). Although accuracy sometimes does increase across repeated attempts, improvement cannot be taken for granted (Johns & Darley, 1970). In his presentation at the 1968 ASHA convention, Darley explained "The patient is often aware of his error but is frequently unable to correct it." As an example, when repeating the word "thickening," one of our participants in the AOS group (who had minimal aphasia) said, "sl…sli…sli..uh…slickening." Self-correction strings are sometimes more diverse, involving both lexical and phonemic/phonetic levels (Goodglass, 1993; Joanette et al., 1980; Marshall & Tompkins, 1982; Lee et al., 2000), as exemplified by another study participant who had AOS with moderate Broca's aphasia and produced the following response when the examiner held up a screwdriver and asked what it is called: "hamber, or…ru…hamber or ru... need to… rrr...fru… I… I know but...sss...skewdiver… screw…screw."

Self-corrections of speech sound errors are not unique to AOS. Many people with aphasia and phonemic paraphasia also attempt to correct their errors. Specific to conduction aphasia, the French term *conduit d'approche* expresses self-correction toward increasing phonologic accuracy (Joanette et al., 1980; Kohn, 1984; Marshall & Tompkins, 1982; Valdois et al., 1989). One of our participants with conduction aphasia struggled with the word "eraser": "es…ester, easer, easer, eiser, hazer, azer, h-a-s-e-r, hasser, easer, hasier, eraser, eraser, pencil eraser." Again, though there may be gradual approximation in the direction of greater accuracy (Gandour et al., 1994; Joanette et al., 1980; Valdois et al., 1989), success is not certain (Farmer et al., 1978; Gandour et al., 1994; Joanette et al., 1980; Kohn, 1989). Of note, the extent to which self-corrections are successful appears to be inversely related to impairment severity (Farmer et al., 1978; Marshall & Tompkins, 1982).

Repeated Production to Self-Correct

The task of repeating words several times consecutively does present at least a theoretical opportunity to self-correct. Small studies and informal observations have examined the accuracy of sequential repetition in AOS, with varied findings. Johns and Darley (1970) asked 10 people with aphasia and AOS to repeat monosyllabic words 3 times successively and reported that the: "…apraxic subjects often seemed to adjust themselves to the task, progressing from incorrect to correct production" (p. 569). In contrast, LaPointe and Horner asked seven other people with aphasia and AOS to repeat multisyllabic

words but requested 10 repetitions (LaPointe & Horner, 1976) and found that accuracy deteriorated across trials. A logical (untested) explanation could be that the monosyllabic words of Johns and Darley (1970) were less error-prone and therefore more readily self-corrected compared to the multisyllabic words of LaPointe and Horner (1976). Neither study compared the performance in these small samples of people with AOS to people with aphasia and no AOS.

Dabul included a "repeated trials test" in the Aphasia Battery for Adults (ABA; Dabul, 1979), using multisyllabic target words that participants were asked to repeat 3 times successively. While field testing the first version of the test, the examiners noted that accuracy improved for some people with AOS from the first to the third repetition but worsened for others, whereas those with aphasia or dysarthria changed little. Based on these observations, Dabul included accuracy directionality across trials in the scoring for both the ABA and the ABA-2. The manuals state that speech accuracy for those with moderate AOS tends to improve across trials, whereas those with severe AOS most often exhibit deterioration across trials. They advise that score changes in either direction should be considered evidence of AOS (Dabul, 1979, 2000).

Until recently, only two studies have compared accuracy on requested successive attempts for AOS relative to aphasia that is uncomplicated by AOS but—to varying degrees—may involve phonologically based sound errors. Both studies were based on the ABA repeated trials subtest (Dabul, 1979, 2000). In the previously mentioned study, McNeil and colleagues found that four speakers with relatively isolated AOS reached accurate production across three repetitions less often than four participants who had conduction aphasia (McNeil et al., 1995). Using the same ABA subtest, but a more diverse and clinically representative sample of people with aphasia with and without AOS, Scholl and colleagues found, on average, that phonemic accuracy deteriorated from the first to the third trial for 20 participants with aphasia and AOS and remained stable for 21 speakers with aphasia only (Scholl et al., 2017). In both studies, the participants with AOS produced far more sound errors than the participants with aphasia only—a complication that is typical of research comparing samples from these populations. It is likely that the degree to which a person can improve production accuracy from trial to trial is determined by the severity of the speech production impairment (or the phonetic complexity of the target words), rather than the presence or absence of an AOS diagnosis. A larger participant sample is needed to evaluate this possibility and reconcile the varied performance patterns that have been observed to date. To this end, this study is based on data from our recent study on error consistency in participants with quantitatively defined articulatory and prosodic difficulties (Haley, Cunningham, Jacks, et al., 2021).

Purpose

We posed two research questions. (a) Does production accuracy improve, deteriorate, or remain the same when people with AOS and people with aphasia are asked to repeat difficult words several times consecutively? (b) If there is a change, does it vary by diagnostic profile or with the severity of overall sound production difficulties?

Method

The study was approved by the institutional review boards of the collaborating universities, and all participants provided signed informed consent.

Participants and Clinical Assessment

We restricted the analysis to those in the original sample who had stroke etiology, produced at least one error on the repeated word task, and gave a response that was sufficiently like the target to allow meaningful analysis of phonetic accuracy and consistency. This resulted in a sample size of 133. Study participants were native speakers of English. Most were from the Southeastern region of the United States; however, there was no effort to document or control regional dialect. All passed a hearing screening at 40 dB HL for 1000 Hz or 2000 Hz in at least one ear (Ventry & Weinstein, 1983). Aphasia severity was estimated with the Western Aphasia Battery–Revised (WAB-R; Kertesz, 2006), yielding a range of Aphasia Quotients (AQs) from 17.0 to 99.6 (Mdn = 73.9, interquartile range [IQR] = 55.2 to 89.2). Fifteen participants scored above the WAB-R cutoff for aphasia (AQ > 93.7); 48 had mild, 44 had moderate, 24 had severe, and two had very severe aphasia per WAB-R criteria (Kertesz, 2006). The first and second authors verified perceptually that no speaker presented with dysarthria of more than mild severity.

Monosyllabic, disyllabic, and multisyllabic word productions from a traditional motor speech evaluation (see Haley et al., 2012) were used to obtain operationalized measures on two primary diagnostic criteria for AOS—abnormal temporal prosody and frequent distortion errors (Ballard et al., 2015; Duffy, 2020). While we recognize that two behavioral metrics can only approximate diagnosis of a multidimensional syndrome, our strategy was to prioritize diagnostic transparency. As an index of prosody, we used the mean duration of syllables in multisyllabic words—the

word syllable duration (WSD; Haley & Jacks, 2019). The estimate of sound distortions was based on narrow phonetic transcription and calculated as the proportion of produced phonemes that were assigned a distinct distortion diacritic mark (e.g., ambiguous voicing, tongue lowering; Haley et al., 2019). Since these two reproducible measures correspond to core diagnostic criteria for AOS (Duffy, 2020; Jacks & Haley, 2021; McNeil et al., 2009; Strand et al., 2014), we considered them suitable proxies for the purpose of the study. Median values (301.5-ms WSD; 9.5% sound distortions) were used to divide the sample into four quadrants. Long syllable duration with high distortion rates was considered indicative of apraxia of speech ("AOS") and identified a sample of 37 participants; normal syllable duration with low distortion rates was considered indicative of aphasia without AOS ("APH") and resulted in another sample of 37 participants. Finally, we formed two "borderline" diagnostic categories that included 30 and 29 participants, respectively: long syllable duration and low distortion rates ("BORD1") and normal syllable duration and high distortion rate ("BORD2"). Speech diagnosis in these borderline groups was equivocal based on our operational criteria. Refer to Table 1 for a summary of participant demographics and clinical test results in each group.

To estimate the broader speech production accuracy, phonetic transcriptions of the single words from the motor speech evaluation were analyzed at a phonemic level. The proportion of target speech sounds produced without a phonemic error was calculated as 100 minus the phonemic edit distance ratio expressed as a percentage. This measure, "motor speech accuracy," corresponds to the percentage of target phonemes produced correctly (Smith et al., 2019). As shown in Table 1, the percentage was lower for participants with a diagnosis of AOS ($M = 79.4\%$) than for those with APH (aphasia only; $M = 91.5\%$)—the same difference that has, so far, been noted in virtually every group comparison involving these diagnostic groups. Accuracy was 86.2% and 86.3% for the BORD1 and BORD2 groups, respectively.

Elicitation and Analysis of Repeated Word Productions

In addition to the portions of the motor speech examination described so far, each participant produced five sequential repetitions of the following words: "artillery," "catastrophe," "impossibility," and "rhinoceros." Like the rest of the motor speech examination, these words were not selected specifically for this study; rather, they were included based on traditional clinical intuition and expertise (Wertz et al., 1984). The experimenter gave a model of each word and requested that the participant repeat it 5 times in a row. The resulting productions (20 for each participant) were the experimental targets for our recent study (Haley, Cunningham, Jacks, et al., 2021) and are further analyzed in the present report.

As described in the original study, two phonetically trained observers completed the transcriptions using computer-readable phonetic characters (Vitevitch & Luce, 2004). Transcription reliability was satisfactory (91.9% point-to-point exact interobserver agreement for individual segments, evaluated on data from 28 randomly selected participants). As an index of phonemic accuracy, the phonetic edit distance ratio was calculated to express, for each speaker and trial, the proportion of target speech sounds that was produced without a phonemic error (Smith et al., 2019). In addition, we calculated a percent change measure for target speech sounds produced correctly, where the number of segments produced correctly on the first trial was subtracted from number correct on the fifth (last) trial and divided by the number correct for the first trial (Scholl et al., 2017).

Table 1. Demographics, aphasia test results, and percent phonemic errors on a motor speech examination (motor speech accuracy). Results are presented for each of the diagnostic groups separately.

Variable	Participant group			
	AOS	**BORD1**	**BORD2**	**APH**
Sample size (*n*)	37	30	29	37
Sex (M:F)	18:19	22:8	16:13	33:4
Age (years)	61 (52–69)	60 (50–70)	58 (53–68)	63 (53–70)
Months postonset	35 (18–66)	30 (11–93)	30 (8–66)	33.5 (34.8)
WAB-R AQ /100	70 (49–85)	71 (50–90)	70 (55–88)	83 (62–93)
Motor speech accuracy (%)	83 (75–89)	90 (85–97)	89 (82–92)	96 (89–98)

Note. Values indicate medians, and interquartile ranges are reported in parentheses. Motor speech accuracy is the percentage of target phonemes produced correctly on the single word repetition section of the motor speech evaluation. AOS = profile most consistent with apraxia of speech; BORD1 and BORD 2 = borderline or inconclusive profile; APH = profile most consistent with aphasia without AOS; WAB-R AQ = Western Aphasia Battery–Revised Aphasia Quotient (Kertesz, 2006).

Analysis Plan

Percentage or proportion measures, such as the phonemic accuracy measure described above, are prone to skewed distributions due to ceiling effects. For this reason, we used a combination of descriptive and nonparametric group comparisons to evaluate accuracy and percent change in segmental accuracy.

For the percent change measure, we used Wilcoxon signed-rank tests to determine if there was significant increase or decrease relative to 0. Because any change relative to 0 might have been positive or negative, we employed two-tailed significance tests. The Wilcoxon tests were followed up with two Kruskal–Wallis analyses to determine whether there was a difference in the magnitude of the change score across groups. The independent variable for the first analysis was diagnostic group, and the independent variable for the second analysis was a motor speech accuracy classification that was formed based on quartiles of overall accuracy from the motor speech examination. The rationale for using motor speech accuracy groups rather than a continuous variable was that the accuracy percentage from the motor speech exam was negatively skewed due to ceiling effects (i.e., a disproportionately high quantity of high accuracy scores). Follow-up comparisons using Wilcoxon tests were used when appropriate.

Results

The serial repetition task was presented the same way to all participants, but not everyone attempted all targets and all repetitions. One hundred twelve participants (84%) produced all 20 words (i.e., five repetitions of all four target words), but 21 produced an average of 15 words because they declined to attempt all trials (two of these repeated only four or five words). All participants were included in the analysis because they completed the serial production task. Median accuracy across trials and participants was 84.6% and varied as expected across groups (AOS: *Mdn* = 70.0%, IQR = 53.8% to 90.9%; BORD1: *Mdn* = 84.6%, IQR = 66.7% to 100.0%; BORD2: *Mdn* = 81.8%, IQR = 60.0% to 100.0%; APH: *Mdn* = 100.0%, IQR = 76.9% to 100.0%).

Our first research question was whether phonemic accuracy, on average, changes across repeated utterances for people with speech production difficulties after left hemisphere stroke. To address this question, Figure 1 shows median production accuracy across repetitions for all participants by diagnosis group, with IQRs reflected in the "box" borders. As illustrated, production accuracy decreased visibly across repetitions for three of the four groups, with only the APH group showing minimal change.

Figure 1. Median phonemic accuracy by repetition trial for the four participant groups. Interquartile range is displayed by the boxes; median value is notched for each box. Whiskers represent minimum and maximum values, excluding outliers. AOS = profile most consistent with apraxia of speech; APH = profile most consistent with aphasia without AOS; BORD1 and BORD 2 = borderline or inconclusive profile.
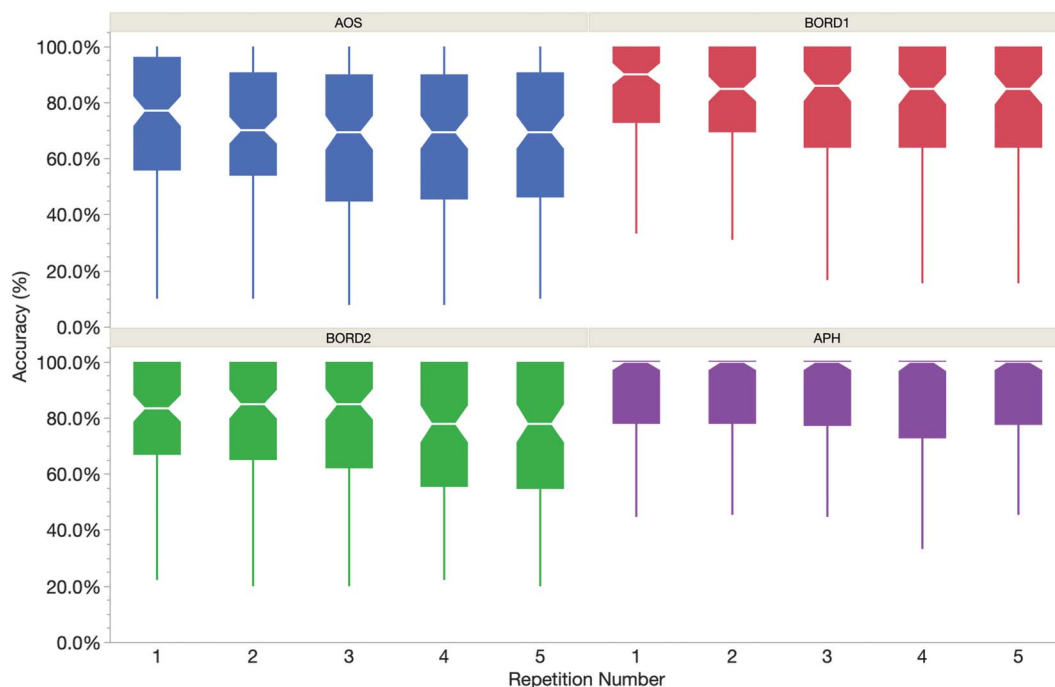
To quantify the magnitude of the change, we examined the median percentage change in correct phonemes from the first to the last production attempt across words. For this analysis, a positive change score indicates that the proportion of phonemes produced correctly increased from the first to the last repetition; a negative change score indicates a decline. Across participants, the median change from first to last production was −4.8% (IQR = −14.3% to 0%). A signed-rank test showed that this change was significantly less than 0 (W = −2,667, p < .0001). The proportion of participants whose accuracy was lower at the last production than at the first production was 61.9%. Only 21.6% improved accuracy across repetitions, and 16.4% showed no numerical change. Thus, we conclude that overall phoneme accuracy worsened from the first to the last attempt in our sample of people with AOS and people with aphasia.

Having demonstrated that production accuracy generally deteriorated across repetitions in the overall sample, we turn to our second question: Does change across repetitions vary by diagnostic profile or with the severity of overall sound production difficulties? While we observed that productions for 61.9% of all participants worsened across repetitions, this proportion differed somewhat across groups, with 67.6% of the AOS declining compared to 60.0% for BORD1, 65.5% for BORD2, and 55.3% for the APH group. Change in accurate segments from first to last production was generally negative, with a median change of −10.0% for the AOS group (IQR = −15.8% to 1.2%), −3.0% for BORD1 (IQR = −16.8% to 0%), −11.1% for BORD2 (IQR = −18.9% to 1.2%), and − 2.3% for the APH group (IQR = −10.8% to 1.2%). Wilcoxon signed-rank tests showed that percent change in accuracy was significantly lower than 0 for all groups (W_{AOS} = −217.0, p = .0005; W_{BORD1} = −145.5, p = .0013; W_{BORD2} = −152.5, p = .0003; W_{APH} = −155, p = .017), and a Kruskal–Wallis rank sum test showed that there was no significant difference in the change magnitude among the four diagnostic groups, $\chi^2(3)$ = 3.01, p = .39.

To determine whether overall speech production accuracy impacted change across productions, we examined data across the four phonemic accuracy groups. The median change varied somewhat by group, with participants in lower accuracy groups showing numerically greater decline across repetitions compared to higher accuracy groups. The lowest accuracy group declined by a median of −11.1% (< 82%, IQR = −18.2% to 0%), with −9.7 for the second quartile (82% to 90% accuracy group, IQR = −15.1% to 2.5%), −3.9% for the third quartile (90% to 95% accuracy, IQR = −12.8% to 0%), and −2.3% for the highest accuracy group (> 95%, IQR = −4.8% to 1.8%). Wilcoxon signed-rank tests showed that the percent change was significantly lower than 0 for all

groups but the highest accuracy group (W_{Q1} = −191.5, p = .0002; W_{Q2} = −242.0, p = .0002; W_{Q3} = −146.5, p = .0012; W_{Q4} = −92.5, p = .08). Although these results seem to indicate an effect of accuracy group on percent change, a Kruskal–Wallis test indicated that group difference was not significant, $\chi^2(3)$ = 5.81, p = .12. Because of the marginal significance of the Kruskal–Wallis test, follow-up Wilcoxon tests were completed, revealing only one significant comparison: a difference between the most accurate and least accurate groups (Z = 2.21, p = .027). We conclude, therefore, that participants with aphasia worsen across trials, regardless of whether they also have AOS and regardless of speech production severity.

Discussion

We recently demonstrated in a large and clinically representative sample that sound production varies from trial to trial in people with aphasia, regardless of whether these participants also show a speech profile consistent with the behavioral definition of AOS (Haley, Cunningham, Jacks, et al., 2021). In that study, consistency of production was not explained by diagnostic group, leading us to conclude that inconsistency is not useful for differential diagnosis between the two behaviorally defined syndromes. The purpose of this study was to determine whether the pattern of change across multiple repetitions predicts differential diagnosis. Results show that it does not. A small, but robust, worsening was evident in the entire sample, regardless of diagnosis. Next, we consider potential explanations for the deterioration in accuracy and whether our results, and the repeated words task itself, might inform clinical evaluation and treatment planning.

Our main question was whether phonemic accuracy worsened across attempts. Results revealed that speech accuracy declined by 4.8% from the first to the last trial, and for the AOS group specifically, the decline was 10.0%. This worsening replicates the early study by LaPointe and Horner (1976), where deterioration occurred across 10 repetitions of multisyllabic words. It should be noted that the participants in the LaPointe and Horner sample were characterized as having aphasia with AOS but selected based on criteria that have since been determined to not differentiate between AOS and phonemic paraphasia ("predominantly substitution errors, variability in error pattern on repeated trials of polysyllabic words, and/or difficulty in the initiation of speech," p. 262). The results are also consistent with those of the four participants with relatively isolated AOS in McNeil et al. (1995) and the 20 participants with aphasia and AOS in Scholl et al. (2017). In these latter two studies, phonemic accuracy was more stable for people diagnosed with aphasia

without AOS. Visual inspection of Figure 1 indicates a similar relationship in our study and suggests that it is explained by ceiling effect where those with aphasia only produce very few errors. Still, most participants in all four diagnostic groups worsened from the first to the last trial, including 68% for the AOS group and 55% for the aphasia-only group. Comparable proportions in the Scholl et al. (2017) study indicated worsening for 80% of participants in the AOS group and 24% of participants in the aphasia group. In this study, the median phonemic accuracy across all segments produced on the repeated word task was 71% for participants with likely AOS and 92% for participants with aphasia, which corresponds to combined substitution, omission, and addition errors at 29% for the AOS group versus 8% for the aphasia group. Per Table 2 in the Scholl et al. study, median percentages for their AOS group were 16% substitutions, 5% omissions, and 3% additions, whereas median percentages for their aphasia-only group were only 2% substitutions, 1% omissions, and 1% additions. Of note, severity differences across groups were also apparent in the study by McNeil and colleagues, though it is unclear to what extent their published summaries correspond with our definitions of phonemic accuracy.

An Articulatory and Cognitive Challenge

In considering why production would worsen over successive repetitions, it is probably important that the serial repetition task is unfamiliar to most people with aphasia and AOS and sufficiently challenging to assume the functional properties of a maximum performance task. Repeated production of the same utterance multiple times is known to trigger semantic satiation, which is the subjective experience that words and utterances temporarily lose some of their meaning, as the talker's attention shifts to phonology (Severance & Washburn, 1907). Salience of word production is reduced and vulnerability to error is particularly high for later trials in the sequence.

The immediate succession of phonetically challenging words requires that the speaker navigates recurring syllable sequences under time pressure. Similarities to tongue-twister phrases and research designs cannot be ignored. Tongue twisters are maximal performance challenges designed to be error inducing for neurotypical speakers. They are assembled in such a way that similar, but distinct, articulatory gestures are juxtaposed and alternating, often through alliteration and articulatory variation involving a single phonetic feature (e.g., between the alveolar and palatal places of articulation for the fricatives in "she sells seashells by the seashore"). When producing a tongue-twister sequence, speakers must rapidly suppress recently activated and similar articulatory gestures as they transition swiftly from one gesture to another. When this fails, there is evidence of interference across words (e.g., "sea" is produced as "she" or with a distortion error that includes traces of both the alveolar and palatal fricatives; Frisch & Wright, 2002; Goldrick & Blumstein, 2006). Timely suppression of recently activated motor programs is particularly challenging when targets are unpracticed and produced under time pressure. Attention to prosody and semantics typically helps the speaker achieve success, but this strategy is not readily available during semantic satiation. Neurotypical speakers are thought to experience a breakdown in executive control when tongue-twister difficulty or rate exceeds their motor programming ability (Shen & Janse, 2020). It is reasonable to expect that speakers with AOS and other postlexical impairments would experience similar difficulties and predictable that their threshold for breakdown would be lower than for neurotypical participants. Therefore, a "tongue twister" effect would be evident for target sequences that normally present a manageable challenge. Indeed, diadochokinetic tasks, such as repeating the nonsense syllable string "puh-tuh-kuh" multiple times, are sufficiently challenging to break production accuracy for most people with aphasia and phonemic or speech motor programming difficulties, but manageable for most neurotypical speakers (Duffy, 2020; Haley et al., 2012).

Mailend and colleagues have proposed that people with AOS struggle to retrieve speech motor programs and that their problems manifest as ineffective suppression of recently activated motor programs, which create interference via activation competition. They demonstrated, in a reaction time study, that people with AOS display exaggerated interference relative to neurotypical speakers and those with only aphasia when they, just before producing a target word, are exposed to a visual or auditory word prime that competes phonetically with the target they have prepared to produce (Mailend et al., 2019; Mailend & Maas, 2013). In a subsequent study, the same authors observed interference effects for participants with AOS upon articulating phonetically similar words consecutively (Mailend et al., 2021)—a task that resembles the serial word repetition task in this study.

If interference causes simultaneous activation of competing motor programs, evidence of blending is expected, such that the phonetic output maintains traces of both targets and therefore involves varying degrees of phonetic ambiguity. This type of "distortion error," which listeners may perceive as a phonemic error (or no error at all), is indeed reported in both acoustic and more fine-grained perceptual studies of AOS and—to a lesser extent—in aphasia (Blumstein et al., 1980; Buchwald & Miozzo, 2012; Hagedorn et al., 2017; Haley, 2002; Haley et al., 2019; Pouplier & Hardcastle, 2005).

While exaggerated in AOS, there is evidence that interference from competing motor programs is present also in neurotypical speakers. Wilshire (1999) asked neurologically healthy participants to repeat strings of four phonetically similar monosyllabic words (e.g., *moss, knife, noose, muff*) 4 times consecutively without pausing and at a rate of 100 words per minute (Wilshire, 1999). The experiment was characterized as a tongue-twister task, and its intended effect was to induce phonemic/phonetic errors. Interestingly, results showed that accuracy deteriorated from the first production to subsequent repetitions, which the author interpreted as decay of a temporary motor plan. Whatever the mechanism may be, there is converging evidence that rapid iterative production of challenging word or syllable strings can induce errors in people with both normal and impaired motor planning for speech, the main difference being the level at which breakdown occurs and errors emerge. It is safe to assume that people with AOS or aphasia with phonemic paraphasia find it challenging to repeat multisyllabic (real and nonsense) words multiple times, particularly when they are produced with phonetically complex segments and stress patterns. These properties introduce challenge in a manner that resembles what a more complex tongue twister does for neurotypical speaker and presents a form of maximum performance challenge.

The Cognition and Psychology of Self-Correction

So far, we have concluded that stroke survivors with aphasia, with and without AOS, are not effectively self-correcting their productions during the commonly administered serial word repetition task. Results aside, is it even realistic to consider the task to be an invitation to self-correct? A person's ability and willingness to recognize and correct speech errors is undoubtedly influenced by a multitude of powerful factors. Better attention to and understanding of these factors may improve the effectiveness of treatment programs (Wambaugh et al., 2016). For example, if a client is working with a speech-language pathologist and produces a straightforward error while the two are working together on accuracy, the client may initiate self-correction promptly and successfully. If, on the other hand, the error has several dimensions and the task is difficult or complex, the client may not be able or willing to recruit cognitive resources to self-evaluate and improve in a second attempt. Motivational factors are equally important because memories of past failures and successes influence whether a person believes a repeated attempt will be successful and whether they deem it worthwhile to try again (Bandura, 1997).

People with aphasia and AOS have personal strategies and hypotheses about why they are asked to perform various tasks in assessment and treatment. It is informative to consider these perspectives and not assume clients perceive the point of the task the same way clinicians do. It can be helpful to simply ask about the strategies they use and how they think task variations affect their performance. One of our participants with aphasia and AOS had completed many previous motor speech evaluations. Concerning the repeated word task, he made the following unsolicited observation: "One word…almost… but two three time—every time mistake." His introspection characterized the performance accurately and seems to validate subjectively the conclusions of this study.

There are other reasons the word repetition task is poorly suited to evaluate self-correction. In our experience, repetition tasks induce few attempts to revise and self-correct (Haley et al., 2012). People usually leave their errors alone, and this is often in stark contrast to how they speak spontaneously or during a naming assessment. Accordingly, two of the three self-correction examples we offered in the introduction were from the WAB-R Naming subtest. Clients' response strategies are likely to vary based on how they understand the purpose of various assessment and treatment tasks. Accordingly, error frequencies would probably have been higher in this study, had we imposed further restriction on rate, for example, through modeling, metronome, or a request to repeat the sequence fast and evenly (as is typical for diadochokinetic maximum performance tasks). Conversely, a different response strategy may have resulted simply by altering the instruction from, "Say this word. Say it five times," to, "Say this word. You have five chances to say it."

Self-Corrections and Motor Learning

In the manual for the ABA, Dabul suggested that the repeated trials test might predict responsiveness to different treatment strategies, so that a client who improves with repetition should be given ample time to self-correct, whereas a patient whose accuracy deteriorates should be stopped and cued toward successful production (Dabul, 1979, 2000). At the time, no research had been conducted to support the hypothesis, and to our knowledge, this remains the case. Observing that performance deteriorates during a maximal performance task does not necessarily indicate poor prognosis for repeated practice. To the contrary, most people with aphasia and AOS do learn and improve over time even when they make mistakes (as neurotypical people also do with tongue twisters). For example, Deal (1974) examined oral reading of a passage 5 times in succession in five participants with AOS. Although the participants tended to make errors on the same words, they improved in overall accuracy across the repeated readings. There is also emerging evidence that

learning occurs in AOS without feedback or cueing, simply from saying words multiple times (Wambaugh et al., 2012).

Many AOS treatments incorporate a request to pause and repeat a successful production several times with the idea that this strengthens learning (Rosenbek et al., 1973; Wambaugh et al., 1998, 2014). It is usually considered essential that the client be guided when to repeat to avoid unintentionally practicing and strengthening erroneous production. The results of this study affirm the need to observe how the client approaches the repetition task and guide them so as to avoid semantic satiation, motor program interference, and unintended worsening. We have observed that people with aphasia and AOS choose, on their own initiative during treatment, to repeat some target words multiple times (Haley, Cunningham, Kim, & Shafer, 2021). By studying how people with aphasia and AOS choose to practice relative to the level of difficulty they are working with, it may be possible to better understand how they learn and whether self-initiated corrections and repetitions can play a functional role in their recovery. There is no reason to assume that their response when asked to say difficult words multiple times would predict this ability or preference.

Conclusions

By examining a large sample of stroke survivors who were assessed and diagnosed with the same procedures, we were able to show that when people with aphasia with and without AOS are asked to repeat challenging words several times, their accuracy varies from trial to trial and, declines from first to last repetition. To explain the deterioration, we conceptualized the request as a maximum performance task and compared it to tongue-twister paradigms that have been used to elicit speech sound errors in neurotypical speakers. We found that the task, which is included in most motor speech examinations, does not help differentiate between AOS and aphasia with phonemic paraphasia. After considering its potential as an index of ability to self-correct, we rejected the idea based on recent studies and suggested, instead, that its primary value is simply to elicit speech errors, which would appear to be particularly informative for people with mild speech production impairments. We recommend clarifying the purpose of the task to diagnosticians and updating assessment materials and procedures accordingly.

Data Availability Statement

Data are available on request from the first and second authors.

References

Bailey, D. J., Blomgren, M., DeLong, C., Berggren, K., & Wambaugh, J. L. (2017). Quantification and systematic characterization of stuttering-like disfluencies in acquired apraxia of speech. *American Journal of Speech-Language Pathology, 26*(2S), 641–648. https://doi.org/10.1044/2017_AJSLP-16-0108

Ballard, K. J., Wambaugh, J. L., Duffy, J. R., Layfield, C., Maas, E., Mauszycki, S., & McNeil, M. R. (2015). Treatment for acquired apraxia of speech: A systematic review of intervention research between 2004 and 2012. *American Journal of Speech-Language Pathology, 24*(2), 316–337. https://doi.org/10.1044/2015_AJSLP-14-0118

Bandura, A. (1997). *Self-efficacy: The exercise of control.* W. H. Freeman.

Bislick, L., McNeil, M., Spencer, K. A., Yorkston, K., & Kendall, D. L. (2017). The nature of error consistency in individuals with acquired apraxia of speech and aphasia. *American Journal of Speech-Language Pathology, 26*(2S), 611–630. https://doi.org/10.1044/2017_AJSLP-16-0080

Blumstein, S. E., Cooper, W. E., Goodglass, H., Statlender, S., & Gottlieb, J. (1980). Production deficits in aphasia: A voice-onset time analysis. *Brain and Language, 9*(2), 153–170. https://doi.org/10.1016/0093-934X(80)90137-6

Buchwald, A., & Miozzo, M. (2012). Phonological and motor errors in individuals with acquired sound production impairment. *Journal of Speech, Language, and Hearing Research, 55*(5), S1573–S1586. https://doi.org/10.1044/1092-4388(2012/11-0200)

Dabul, B. (1979). *Apraxia Battery for Adults.* C. C. Publications.

Dabul, B. (2000). *Apraxia Battery for Adults-2.* Pro-Ed.

Darley, F. L. (1968). *Apraxia of speech: 107 years of terminological confusion.* Presented at the Annual Meeting of the American Speech and Hearing Association.

Darley, F. L., Aronson, A. E., & Brown, J. R. (1975). *Motor speech disorders.* Saunders.

Deal, J. L. (1974). Consistency and adaptation in apraxia of speech. *Journal of Communication Disorders, 7*(2), 135–140. https://doi.org/10.1016/0021-9924(74)90026-4

Duffy, J. R. (2020). *Motor speech disorders: Substrates, differential diagnosis, and management* (4th ed.). Elsevier.

Farmer, A., O'Connell, P. F., & O'Connell, E. J. (1978). Sound error self-correction in the conversational speech of nonfluent and fluent aphasics. *Folia Phoniatrica, 30*(4), 293–302. https://doi.org/10.1159/000264138

Frisch, S. A., & Wright, R. (2002). The phonetics of phonological speech errors: An acoustic analysis of slips of the tongue.

Journal of Phonetics, 30(2), 139–162. https://doi.org/10.1006/jpho.2002.0176

Gandour, J., Akamanon, C., Dechongkit, S., Khunadorn, F., & Boonklam, R. (1994). Sequences of phonemic approximations in a Thai conduction aphasic. *Brain and Language, 46*(1), 69–95. https://doi.org/10.1006/brln.1994.1005

Goldrick, M., & Blumstein, S. E. (2006). Cascading activation from phonological planning to articulatory processes: Evidence from tongue twisters. *Language and Cognitive Processes, 21*(6), 649–683. https://doi.org/10.1080/01690960500181332

Goodglass, H. (1993). *Understanding aphasia*. Academic Press.

Hagedorn, C., Proctor, M., Goldstein, L., Wilson, S. M., Miller, B., Gorno-Tempini, M. L., & Narayanan, S. S. (2017). Characterizing articulation in apraxic speech using real-time magnetic resonance imaging. *Journal of Speech, Language, and Hearing Research, 60*(4), 877–891. https://doi.org/10.1044/2016_JSLHR-S-15-0112

Haley, K. L. (2002). Temporal and spectral properties of voiceless fricatives in aphasia and apraxia of speech. *Aphasiology, 16*(4–6), 595–607. https://doi.org/10.1080/02687030244000257

Haley, K. L., Cunningham, K. T., Eaton, C. T., & Jacks, A. (2018). Error consistency in acquired apraxia of speech with aphasia: Effects of the analysis unit. *Journal of Speech, Language, and Hearing Research, 61*(2), 210–226. https://doi.org/10.1044/2017_JSLHR-S-16-0381

Haley, K. L., Cunningham, K. T., Jacks, A., Richardson, J. D., Harmon, T., & Turkeltaub, P. E. (2021). Repeated word production is inconsistent in both aphasia and apraxia of speech. *Aphasiology, 35*(4), 518–538. https://doi.org/10.1080/02687038.2020.1727837

Haley, K. L., Cunningham, K. T., Kim, I., & Shafer, J. S. (2021). Autonomy-supportive treatment for acquired apraxia of speech: Feasibility and therapeutic effect. *Aphasiology, 35*(4), 539–559. https://doi.org/10.1080/02687038.2019.1705662

Haley, K. L., & Jacks, A. (2019). Word-level prosodic measures and the differential diagnosis of apraxia of speech. *Clinical Linguistics & Phonetics, 33*(5), 479–495. https://doi.org/10.1080/02699206.2018.1550813

Haley, K. L., Jacks, A., & Cunningham, K. T. (2013). Error variability and the differentiation between apraxia of speech and aphasia with phonemic paraphasia. *Journal of Speech, Language, and Hearing Research, 56*(3), 891–905. https://doi.org/10.1044/1092-4388(2012/12-0161)

Haley, K. L., Jacks, A., de Riesthal, M., Abou-Khalil, R., & Roth, H. L. (2012). Toward a quantitative basis for assessment and diagnosis of apraxia of speech. *Journal of Speech, Language, and Hearing Research, 55*(5), S1502–S1517. https://doi.org/10.1044/1092-4388(2012/11-0318)

Haley, K. L., & Martin, G. (2011). Production variability and single word intelligibility in aphasia and apraxia of speech. *Journal of Communication Disorders, 44*(1), 103–115. https://doi.org/10.1016/j.jcomdis.2010.07.007

Haley, K. L., Smith, M., & Wambaugh, J. L. (2019). Sound distortion errors in aphasia with apraxia of speech. *American Journal of Speech-Language Pathology, 28*(1), 121–135. https://doi.org/10.1044/2018_AJSLP-17-0186

Harmon, T. G., Jacks, A., & Haley, K. L. (2019). Speech fluency in acquired apraxia of speech during narrative discourse: Group comparisons and dual-task effects. *American Journal of Speech-Language Pathology, 28*(2S), 905–914. https://doi.org/10.1044/2018_AJSLP-MSC18-18-0107

Jacks, A., & Haley, K. (2021). Apraxia of speech. In J. Damico, N. Muller, & M. J. Ball (Eds.), *The handbook of language speech disorders* (2nd ed., pp. 368–390). Wiley-Blackwell. https://doi.org/10.1002/9781119606987.ch17

Joanette, Y., Keller, E., & Lecours, A. R. (1980). Sequences of phonemic approximations in aphasia. *Brain and Language, 11*(1), 30–44. https://doi.org/10.1016/0093-934x(80)90107-8

Johns, D. F., & Darley, F. L. (1970). Phonemic variability in apraxia of speech. *Journal of Speech and Hearing Research, 13*(3), 556–583. https://doi.org/10.1044/jshr.1303.556

Kertesz, A. (2006). *Western Aphasia Battery–Revised*. Pearson.

Kohn, S. E. (1984). The nature of the phonological disorder in conduction aphasia. *Brain and Language, 23*(1), 97–115. https://doi.org/10.1016/0093-934x(84)90009-9

Kohn, S. E. (1989). The nature of the phonemic string deficit in conduction aphasia. *Aphasiology, 3*(3), 209–239. https://doi.org/10.1080/02687038908248992

LaPointe, L. L., & Horner, J. (1976). Repeated trials of words by patients with neurogenic phonological selection-sequencing impairment (apraxia of speech). *Clinical Aphasiology, 6*, 261–277.

Lee, R. K.-S., Yiu, E. M.-L., & Stonham, J. (2000). Phonological disruption and subsequent self-correcting behaviour in Cantonese aphasic speakers. *International Journal of Language & Communication Disorders, 35*(4), 475–486. https://doi.org/10.1080/136828200750001241

Liss, J. M. (1998). Error-revision in the spontaneous speech of apraxic speakers. *Brain and Language, 62*(3), 342–360. https://doi.org/10.1006/brln.1997.1907

Mailend, M.-L., & Maas, E. (2013). Speech motor programming in apraxia of speech: Evidence from a delayed picture-word interference task. *American Journal of Speech-Language Pathology, 22*(2), S380–S396. https://doi.org/10.1044/1058-0360(2013/12-0101)

Mailend, M.-L., Maas, E., Beeson, P. M., Story, B. H., & Forster, K. I. (2019). Speech motor planning in the context of phonetically similar words: Evidence from apraxia of speech and aphasia. *Neuropsychologia, 127*, 171–184. https://doi.org/10.1016/j.neuropsychologia.2019.02.018

Mailend, M.-L., Maas, E., Beeson, P. M., Story, B. H., & Forster, K. I. (2021). Examining speech motor planning difficulties in apraxia of speech and aphasia via the sequential production of phonetically similar words. *Cognitive Neuropsychology, 38*(1), 72–87. https://doi.org/10.1080/02643294.2020.1847059

Marshall, R. C., & Tompkins, C. A. (1982). Verbal self-correction behaviors of fluent and nonfluent aphasic subjects. *Brain and Language, 15*(2), 292–306. https://doi.org/10.1016/0093-934x(82)90061-x

Mauszycki, S. C., Wambaugh, J. L., & Cameron, R. M. (2010). Variability in apraxia of speech: Perceptual analysis of monosyllabic word productions across repeated sampling times. *Aphasiology, 24*(6–8), 838–855. https://doi.org/10.1080/02687030903438516

McNeil, M. R., Odell, K. H., Miller, S. B., & Hunter, L. (1995). Consistency, variability, and target approximation for successive speech repetitions among apraxic, conduction aphasic, and ataxic dysarthric speakers. *Clinical Aphasiology, 23*, 39–55.

McNeil, M. R., Robin, D. A., & Schmidt, R. A. (2009). Apraxia of speech: Definition, differentiation, and treatment. In M. R. NcNeil (Ed.), *Clinical management of sensorimotor speech disorders* (2nd ed., pp. 249–268). Thieme.

Molloy, J., & Jagoe, C. (2019). Use of diverse diagnostic criteria for acquired apraxia of speech: A scoping review. *International Journal of Language & Communication Disorders, 54*(6), 875–893. https://doi.org/10.1111/1460-6984.12494

Odell, K., McNeil, M. R., Rosenbek, J. C., & Hunter, L. (1991). Perceptual characteristics of vowel and prosody production in apraxic, aphasic, and dysarthric speakers. *Journal of Speech and Hearing Research, 34*(1), 67–80. https://doi.org/10.1044/jshr.3401.67

Pouplier, M., & Hardcastle, W. (2005). A re-evaluation of the nature of speech errors in normal and disordered speakers. *Phonetica, 62*(2–4), 227–243. https://doi.org/10.1159/000090100

Rosenbek, J. C., Lemme, M. L., Ahern, M. B., Harris, E. H., & Wertz, R. T. (1973). A treatment for apraxia of speech in adults. *Journal of Speech and Hearing Disorders, 38*(4), 462–472. https://doi.org/10.1044/jshd.3804.462

Scholl, D. I., McCabe, P. J., Heard, R., & Ballard, K. J. (2017). Segmental and prosodic variability on repeated polysyllabic word production in acquired apraxia of speech plus aphasia. *Aphasiology, 32*(5), 578–597. https://doi.org/10.1080/02687038.2017.1381876

Severance, E., & Washburn, M. F. (1907). The loss of associative power in words after long fixation. *American Journal of Psychology, 18*(2), 182. https://doi.org/10.2307/1412411

Shen, C., & Janse, E. (2020). Maximum speech performance and executive control in young adult speakers. *Journal of Speech, Language, and Hearing Research, 63*(11), 3611–3627. https://doi.org/10.1044/2020_JSLHR-19-00257

Shuster, L. I., & Wambaugh, J. L. (2008). Token-to-token variability in adult apraxia of speech: A perceptual analysis. *Aphasiology, 22*(6), 655–669. https://doi.org/10.1080/02687030701632161

Smith, M., Cunningham, K. T., & Haley, K. L. (2019). Automating error frequency analysis via the phonemic edit distance ratio. *Journal of Speech, Language, and Hearing Research, 62*(6), 1719–1723. https://doi.org/10.1044/2019_JSLHR-S-18-0423

Staiger, A., & Ziegler, W. (2008). Syllable frequency and syllable structure in the spontaneous speech production of patients with apraxia of speech. *Aphasiology, 22*(11), 1201–1215. https://doi.org/10.1080/02687030701820584

Strand, E. A., Duffy, J. R., Clark, H. M., & Josephs, K. (2014). The apraxia of speech rating scale: A tool for diagnosis and description of apraxia of speech. *Journal of Communication Disorders, 51,* 43–50. https://doi.org/10.1016/j.jcomdis.2014.06.008

Valdois, S., Joanette, Y., & Nespoulous, J.-L. (1989). Intrinsic organization of sequences of phonemic approximations: A preliminary study. *Aphasiology, 3*(1), 55–73. https://doi.org/10.1080/02687038908248976

Ventry, I., & Weinstein, B. (1983). Identification of elderly people with hearing problems. *ASHA, 25*(7), 37–42.

Vitevitch, M. S., & Luce, P. A. (2004). A web-based interface to calculate phonotactic probability for words and nonwords in English. *Behavior Research Methods, Instruments, & Computers, 36*(3), 481–487. https://doi.org/10.3758/BF03195594

Wambaugh, J. L., Duffy, J. R., McNeil, M. R., Robin, D. A., & Rogers, M. A. (2006). Treatment guidelines for acquired apraxia of speech: A synthesis and evaluation of the evidence. *Journal of Medical Speech-Language Pathology, 14*(2), 15–34.

Wambaugh, J. L., Kalinyak-Fliszar, M. M., West, J. E., & Doyle, P. J. (1998). Effects of treatment for sound errors in apraxia of speech and aphasia. *Journal of Speech, Language, and Hearing Research, 41*(4), 725–743. https://doi.org/10.1044/jslhr.4104.725

Wambaugh, J. L., Nessler, C., Cameron, R., & Mauszycki, S. C. (2012). Acquired apraxia of speech: The effects of repeated practice and rate/rhythm control treatments on sound production accuracy. *American Journal of Speech-Language Pathology, 21*(2), S5–S27. https://doi.org/10.1044/1058-0360(2011/11-0102)

Wambaugh, J. L., Shuster, L., Bailey, D. J., Mauszycki, S., Kean, J., Nessler, C., Wright, S., & Brunsvold, J. (2016). Self-judgments of word production accuracy in acquired apraxia of speech. *American Journal of Speech-Language Pathology, 25*(4S), S716–S728. https://doi.org/10.1044/2016_AJSLP-15-0139

Wambaugh, J. L., Wright, S., Nessler, C., & Mauszycki, S. C. (2014). Combined Aphasia and Apraxia of Speech Treatment (CAAST): Effects of a novel therapy. *Journal of Speech, Language, and Hearing Research, 57*(6), 2191–2207. https://doi.org/10.1044/2014_JSLHR-L-14-0004

Wertz, R. T., LaPointe, L. L., & Rosenbek, J. C. (1984). *Apraxia of speech in adults: The disorder and its management.* Grune & Stratton.

Wilshire, C. E. (1999). The "tongue twister" paradigm as a technique for studying phonological encoding. *Language and Speech, 42*(1), 57–82. https://doi.org/10.1177/00238309990420010301