

# **The Nature of Error Consistency in Individuals with Apraxia of Speech and Aphasia**

Lauren Bislick<sup>1,2</sup>

Malcolm McNeil<sup>3,4</sup>

Kristie A. Spencer<sup>1</sup>

Kathryn Yorkston<sup>1,5</sup>

Diane L. Kendall<sup>1,2,6</sup>

<sup>1</sup>Department of Speech and Hearing Sciences, University of Washington, Seattle , Washington

<sup>2</sup>Veterans Administration Puget Sound Health Care System, Seattle , Washington

<sup>3</sup>Department of Communication Sciences and Disorders, University of Pittsburgh, Pennsylvania

<sup>4</sup>Veterans Administration Pittsburgh Health Care System, Pennsylvania

<sup>5</sup>Department of Rehabilitation Medicine, University of Washington, Seattle , Washington

<sup>6</sup>Department of Speech-Language Pathology and Audiology, University of Pretoria, Gauteng

South Africa

Correspondence to Lauren Bislick: [lbislick@uw.edu](mailto:lbislick@uw.edu)

Department of Speech and Hearing Sciences,

University of Washington

Eagleson Hall

1417 N.E. 42nd Street

Box 354875

Seattle, WA 98105-6246

## **Abstract**

**Purpose:** The primary characteristics used to define acquired apraxia of speech (AOS) have evolved to better reflect a disorder of motor planning/programming. However, there is debate regarding the feature of relatively consistent error location and type.

**Method:** Ten individuals with acquired AOS and aphasia and 11 individuals with aphasia without AOS participated in this study. In the context of a two-group experimental design, error consistency was examined via five repetitions of 30 multisyllabic words. The influence of error rate, severity of impairment, and stimulus presentation condition (blocked vs random) on error consistency was also explored, as well as between-group differences in the types of errors produced.

**Results:** Groups performed similarly on consistency of error location; however, individuals with AOS demonstrated greater variability of error type in a blocked presentation condition only. Stimulus presentation condition, error rate, and severity of impairment did not influence error consistency in either group. Lastly, groups differed in the production of phonetic errors (e.g., sound distortions), but not phonemic errors.

**Conclusions:** Overall, findings do not support relatively consistent errors as a differentiating characteristic of AOS.

## Introduction

Acquired apraxia of speech (AOS) is a motor speech disorder that typically results from left hemisphere cerebral vascular accident (CVA), frequently co-occurs with aphasia and dysarthria, and rarely occurs in isolation (Duffy, 2013). The speech characteristics observed in AOS reflect disruption in the ability to plan/program the spatial and temporal movements for fluid speech production (McNeil, Robin, Schmidt, 2009; van der Merve, 2009) and can range in severity from minimal disruption to an inability to produce speech entirely.

The co-occurrence of AOS with aphasia and dysarthria has made it difficult to identify clinical speech characteristics unique to the disorder, and as a result, accurate and reliable diagnosis of the disorder can be challenging (McNeil et al., 2009). In particular, it can be especially difficult to differentiate errors associated with impaired motor planning/programming in AOS from the phonemic paraphasias (PP) observed in aphasia (Duffy, 2013). Over the years, the defining characteristics of AOS have evolved to more sensitively reflect a motor planning/programming deficit, thereby improving diagnostic accuracy. The current discriminatory characteristics of AOS include 1) slow speech rate characterized by prolonged segment and intersegment durations, 2) sound distortions, 3) distorted sound substitutions, 4) relatively consistent errors in regard to error location and error type, and 5) prosodic abnormalities (McNeil et al., 2009; Wambaugh, Duffy, McNeil, Robin & Rogers, 2006). Many of these characteristics have been considered typical of AOS for decades, especially prolonged segment and intersegment durations and prosodic abnormalities (Kent & Rosenbek, 1983; Odell, McNeil, Rosenbek, & Hunter, 1991; Strand & McNeil, 1996). The characteristic of distorted sound errors was first established by investigations of individuals with relatively isolated AOS (Odell, McNeil, Rosenbek, & Hunter, 1990; Square, Darley, & Sommers, 1982) and was later

supported by investigations of individuals with AOS and concomitant aphasia (Haley, Bays, & Ohide, 2001; Mauszycki, Wambaugh, & Cameron, 2010a, 2010b, 2012). The characteristic of error consistency, however, is considered controversial, particularly because of its “relative” description, and has been the topic of more recent debate (Haley, Jacks, Cunningham, 2013; Staiger, Finger-Berg, Aichert, Ziegler, 2012; Ziegler, Aichert, Staiger, 2012).

### ***Error Consistency***

The characteristic of error consistency includes two components: 1) the “consistency” of error location, and 2) the “variability” of error type. Specifically, the consistency of error location refers to the extent to which an error occurs on the same target sound within a word across repeated trials. The variability of error type refers to the extent to which the *same error* is made within the *same location* of a word across repeated trials (McNeil, Odell, Miller, & Hunter, 1995). Initially, inconsistent or variable errors were considered a hallmark of AOS and were used as a clinical marker to aid in the differential diagnosis of AOS from dysarthria (Johns & Darley, 1970). This view, however, was challenged after an influential study revealed that errors in AOS might not be as variable as previously described (McNeil, et al., 1995). In 1995, McNeil and colleagues examined error consistency in individuals with isolated AOS (n = 4) compared to isolated PP (n = 4) during the repetition of two-, three-, and five-syllable words three times consecutively. Participant responses were analyzed via narrow phonetic transcription, which is particularly sensitive to phonetic errors produced by individuals with AOS (Odell et al., 1990). Findings revealed that sound errors produced by individuals with isolated AOS were relatively consistent compared to isolated PP.

Since that time, the characteristic of relative consistency of error location and type has been considered a primary characteristic of AOS and is recommended to help differentially diagnose AOS from aphasia with PP. A recent study aimed to verify the findings of McNeil et al. (1995) in a participant sample more representative of the clinical population served by speech-language pathologists (SLPs). Haley and colleagues (2013) examined error consistency in 32 individuals with aphasia and “segmental speech errors” during the repetition of four- to eight-syllable words five times consecutively. The 32 participants were separated into four groups based on speech characteristics, including prominent speech sound errors and impaired prosody (Group 1; n = 9), speech sound errors and borderline prosody impairment (Group 2; n = 6), sound substitutions errors and normal prosody (Group 3; n = 11), and normal prosody and minimal articulatory errors (Group 4; n = 6). Groups 1 and 2 demonstrated characteristics similar to AOS and concomitant aphasia, whereas Group 3 demonstrated characteristics similar to aphasia with PP, without AOS. Participant responses were transcribed via broad phonetic transcription (vs. narrow). Therefore, although considered a primary characteristic of AOS, phonetic errors such as sound distortions were not captured in this analysis. Results of this study revealed no significant between-group differences in error consistency, conflicting with McNeil and colleagues (1995). Inconsistent findings likely reflect methodological differences, including participant inclusion criteria and sensitivity of the analysis employed.

Other investigations have examined error consistency in individuals with AOS and concomitant aphasia without the use of a comparison group to examine its “relative” nature. Results of these investigations also yield conflicting results. Some studies suggest that errors are consistent (Mauszycki, Dromey, & Wambaugh, 2007; Mauszycki et al., 2010a, 2010b, 2012; Wambaugh, Nessler, Bennett, & Mauszycki, 2004), whereas others indicate that errors are more

variable (Mauszycki & Wambaugh, 2006; Staiger et al., 2012). Methodological differences, such as participant inclusion criteria and the way in which error consistency was defined and examined, likely contribute to inconsistent findings across studies (Shuster & Wambaugh, 2008).

### ***Theoretical Support***

Currently there are no models of speech production that specifically address the nature of error consistency in individuals with AOS or aphasia. Nevertheless, the two models briefly discussed below provide some insight into the issue of error consistency in these populations.

One hypothesis of AOS is that the disorder occurs as a result of disruption to the feedforward commands (i.e., stored motor plans/programs) for speech production. This hypothesis stems from a well-supported framework of speech motor control and a corresponding computational model of speech acquisition and production called the Directions into the Velocities of Speech (DIVA) model (Guenther, 2006; Guenther, Ghosh & Tourville, 2006). The model consists of two control subsystems: feedback control and feedforward control. Feedback control is established first and aids in the development of feedforward control, which consists of learned feedforward commands responsible for the production of rapid speech movements that are relatively consistent from trial to trial. Impaired feedforward control results in inaccurate or “noisy” motor commands and an increased reliance on feedback control (Jacks, 2008; Maas, Mailend & Guenther, 2015). Feedback control is likened to that of speech development, where prior to the construction of feedforward commands, speech production is slower and more variable across repeated trials of the same behavior (Jacks, 2008; Mass et al., 2015). This hypothesized reliance on feedback control predicts that individuals with AOS will demonstrate some degree of variability across trials (i.e., not 100% consistent).

Errors resulting from impaired linguistic processing in aphasia, however, are believed to be more variable than errors resulting from impaired motor planning/programming (Duffy, 2013; McNeil et al., 2009). Dell's Interactive Activation (IA; Dell, 1986) model is a two-step model of lexical retrieval constructed from speech error data that provides some support for variable errors in individuals with aphasia. According to the model, the lexical network consists of three levels of representation: semantic, lexical, and phonemic. Information spreads bi-directionally within the lexical network via connections within and between levels. Errors are proposed to result from interference or "noise" and can occur at any level of the linguistic network. In particular, phonemic paraphasias result from interference that leads to the activation of an incorrect phoneme during phonemic encoding (Dell, 1988). The model attributes interference to three different sources, including spread of activation throughout the linguistic network, previously spoken and upcoming targets, and finally "extraneous cognition and perception" (Dell, 1988). Thus, the potential exists for a number of phonological errors to occur within any given word, especially in an impaired network.

### ***Influential Variables***

The extant literature suggests that error consistency in individuals with AOS may be influenced by a number of variables. Two prominent variables are the motoric complexity of the target sound and the context in which it is presented (Odell et al., 1990). Previous studies suggest that individuals with AOS demonstrate more frequent errors on fricatives compared to stops (Johns & Darley, 1970; LaPointe & Johns, 1975; Odell et al., 1990), clusters compared to singletons (Aichert & Ziegler, 2008; Buchwald & Miozzo, 2012; Odell et al., 1990), trisyllabic compared to monosyllabic words (Marquardt, Schneider, & Jacks, 2010; Strand & McNeil,

1996), and word onsets compared to other locations within a word (Odell et al., 1991; Staiger et al., 2012). In the context of the DIVA model, it is plausible that complex targets increase the motor planning load (van der Merwe, 2009) and further tax the feedback control subsystem, resulting in more reliable and/or predictable error location across multiple trials.

Variables such as stimulus presentation condition (Wambaugh et al., 2004; Stager et al., 2012) error rate (Haley et al., 2013), and severity of impairment (Duffy, 2013; Shuster & Wambaugh, 2008) may also influence performance and should be taken into consideration when examining error consistency. Few studies have examined the effect of stimulus presentation condition – blocked versus random – on word repetition in individuals with AOS and concomitant aphasia. Most often, error consistency is examined via repetition of multisyllabic words, three to five times consecutively, in a blocked presentation condition. Motor learning theory suggests that stimulus presentation condition may affect the way in which motor plans/programs are retrieved (Maas, et al., 2008; Schmidt & Lee, 2005). It is hypothesized that stimuli presented in a blocked condition require the execution of the same motor program across consecutive trials (e.g., BBBB, AAAA, CCCC), whereas stimuli presented in a random presentation condition require the retrieval and/or construction of a different motor program on every trial (e.g., ABAC, BCAC, CABA; Knock et al., 2000). Thus, stimuli presented in a blocked condition may result in more consistent errors, whereas stimuli presented in a random condition may elicit more variable errors. Past investigations, however, have demonstrated inconsistent results in individuals with AOS (Mauszycki et al., 2010a, 2010b, 2012; Johns & Darley, 1970; Wambaugh et al., 2004; Staiger et al., 2012).

With regard to aphasia, the IA model (Dell, 1986) similarly suggests that stimulus presentation condition may influence the retrieval of phonemes for speech production.

Specifically, stimuli presented in a blocked condition require maintained activation of the same phonological representation across consecutive trials, whereas stimuli presented in a random condition require activation of new phonological representations on every trial. The activation of new phonological representations, in the random condition, may provide greater opportunity for interference and, therefore, more variable productions from trial to trial compared to the blocked condition. This phenomenon has not been examined in individuals with aphasia without AOS.

Error rate refers to the frequency in which errors occur within a speech sample. The results of Haley et al. (2013) suggest that error rate may influence measures of error consistency in a group of individuals with aphasia, with and without AOS. Specifically, participants who produced relatively more frequent errors demonstrated higher consistency of error location, but also higher variability of error type compared to participants who produce fewer errors. According to Dell and colleagues, the primary difference between the errors made by normal healthy adults compared to individuals with aphasia is the frequency in which errors occur (Dell, Schwartz, Martin, Saffran & Gagnon, 1997). Thus, differences in error rate likely reflect the effects of brain injury and may be suggestive of the extent of damage or interference in the linguistic network. Similarly, error rate in AOS may reflect the extent of damage to feedforward control and the subsequent reliance on feedback control.

Error rate may be one way in which to measure participant severity in these populations, however, this method has not been extensively examined. Furthermore, the influence of severity of impairment on error consistency is unclear in the literature. For example, Duffy (2013) suggests that more severely impaired individuals with AOS demonstrate more consistent and predictable errors compared to those with milder deficits. Conversely, results of a single-subject

investigation by Shuster & Wambaugh (2008) indicate that errors may be less consistent in more severely impaired speakers with AOS.

The majority of investigations that have examined these influential variables have primarily focused on AOS, thus, even less is known about the influence of these variables on error consistency in aphasia without AOS. Stimulus presentation condition, error rate, and severity of impairment may influence error consistency differently in each of these populations. Furthermore, the heterogeneity often associated with aphasia and AOS, i.e., different impairment profiles or subtypes, warrant further within group examination.

### **Purpose**

The primary objective of this study was to examine the nature of error consistency in individuals with AOS and concomitant aphasia compared to individuals with aphasia without AOS. The influence of stimulus presentation condition, error rate, and severity of impairment on participant performance was also explored, as well as between-group differences in the types of errors produced. In the context of a two-group experimental design, the following primary and secondary research questions were addressed:

Research Question 1. Is there a significant between-group difference in consistency of error location and variability of error type during repetition of two, three, and five syllable words in a blocked presentation condition?

Research Question 2. Is there a significant within-group effect of:

2a) stimulus presentation condition (blocked vs. random) on error consistency outcomes?

2b) error rate, as measured by mean number of sound errors per word, on error consistency outcomes?

2c) severity of impairment, as measured by expert ratings (1=mild, 2=moderate, 3=severe) on error consistency outcomes?

Research Question 3. Is there a significant between-group difference in the types of sound errors produced, as measured by group differences in the production of sound distortions, distorted substitutions, distorted additions, substitutions, omissions, and additions during the repetition of two, three, and five syllable words?

## **Method**

### ***Participants***

Ten individuals with acquired AOS and aphasia (Group A) and 11 individuals with aphasia without AOS (Group P) participated in this study. Group A included four males and six females, ranging in age from 45 to 71 ( $M = 61.8$  years,  $SD = 8.2$  years). Time post onset of stroke ranged from ten months to 230 months ( $M = 82.3$ ,  $SD = 70.8$ ). Group P included five males and six females, ranging in age from 49 to 91 ( $M = 65.6$  years,  $SD = 10.8$  years). Time post onset from stroke ranged from seven months to 125 months ( $M = 67.6$  months,  $SD = 36.2$  months). See Appendix A for individual demographics.

All participants demonstrated aphasia with or without AOS and were at least six months post onset of left hemisphere CVA. Participants also met the following inclusion criteria: (1) right-handed, (2) English as a primary language, (3) minimum of high-school education, (4) passing an audiometric pure-tone, air conduction screening at 35 dB HL at 500, 1K, and 2K Hz for at least one ear (or had reports of adequate hearing when aided;  $n=2$ ), (5) normal or corrected to normal visual acuity (20/20 – 20/40) as determined by a Snellen chart screening, and (6) score

above a 23/36 on the Raven's Coloured Progressive Matrices<sup>1</sup> (RCPM; Raven, Raven, & Court, 1998). Participants were excluded from the study if they had a positive medical history of depression (untreated) or other psychiatric illness, degenerative neurological illnesses, chronic medical illness, or dysarthria. Presence of dysarthria was assessed via physical examination of the jaw, lips, tongue, velopharyngeal function, and respiration and phonation (Yorkston, Beukelman, Strand, & Hakel, 2010) during a structural functional exam and perceptual judgment of speech during the testing session.

Presence of AOS was determined by primary characteristics observed during performance on subtests (I, II, IV, V) of the Apraxia Battery for Adults – Second edition (ABA-2; Dabul, 2000), as well as the picture description task, and the Story Retell Procedure (SRP; McNeil et al., 2007). A diagnosis of AOS was made if the following behaviors were observed: (1) slow speech rate characterized by sound, syllable or word segregation and phoneme lengthening, (2) sound distortions, (3) distorted sound substitutions, and (4) prosodic abnormalities (McNeil et al., 2009; Wambaugh et al., 2006).

Aphasia was determined by the presence of language processing deficits as measured by performance on the language battery portion of the Comprehensive Aphasia Test (CAT; Swinburn, Porter, & Howard, 2004). To further assess presence of phonological deficits in aphasia, the Standardized Assessment of Phonology in Aphasia (SAPA; Kendall et al., 2010) was administered (for a brief description of the SAPA please refer to: Brookshire, Conway, Hunting Pompon, Oelke, & Kendall, 2014). All 21 participants demonstrated aphasia, of varying severity, with phonemic paraphasias (Table 1). All tests and tasks were administered and scored by a certified SLP (the first author).

---

<sup>1</sup> One participant, PP11 was only given one subtest of the RCPM and scored 7/12.

**Table 1. Participant Performance on Measures of Aphasia and Severity Rating**

CAT = Comprehensive Aphasia Test (Swinburn et al., 2004); SAPA = Standardized Assessment of Phonology in Aphasia (Kendall et al., 2010)

	Group A										Group P										AVG (SD)			
	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10		P11		
<b>CAT Subtest</b>	Spoken Language Comprehension (66)	51	53	63	52	41	56	63	59	62	60	<b>56</b> <b>(6.7)</b>	60	22	30	59	47	52	55	51	42	38	29	<b>44.1</b> <b>(12.3)</b>
	Written Language Comprehension (62)	49	53	59	46	44	55	58	58	60	56	<b>53.8</b> <b>(5.4)</b>	58	42	45	44	45	53	50	49	42	33	41	<b>45.6</b> <b>(6.4)</b>
	Repetition (74)	70	34	65	44	59	63	71	22	68	57	<b>55.3</b> <b>(15.8)</b>	64	40	42	61	56	65	46	32	57	26	24	<b>46.6</b> <b>(14.3)</b>
	Naming (58) + Fluency	38+18	56+14	56+16	26+4	28+13	46+11	57+24	18+11	38+12	52+14	<b>55.2</b> <b>(16.8)</b>	52+19	3+3	21+13	48+19	34+18	51+23	40+19	46+20	23+9	4+0	9+3	<b>43.4</b> <b>(25.7)</b>
	Reading (70)	70	29	56	14	44	52	62	58	63	56	<b>49.6</b> <b>(16.9)</b>	66	29	25	67	47	62	55	59	50	0	NA	<b>46</b> <b>(20.5)</b>
	Writing (76)	69	71	73	33	60	67	74	76	71	75	<b>66.9</b> <b>(12.1)</b>	67	24	53	71	56	71	60	64	NA	38	NA	<b>56</b> <b>(14.9)</b>
	Reading (52)	43	20	38	18	28	36	42	28	38	40	<b>33.1</b> <b>(9.0)</b>	47	15	18	43	28	33	35	47	40	2	10	<b>28.9</b> <b>(15.6)</b>
<b>SAPA</b>	Auditory Phonological Processing (65)	55	49	61	47	54	62	56	50	53	55	<b>54.2</b> <b>(4.8)</b>	56	38	39	53	51	60	58	42	59	40	22	<b>47.1</b> <b>(11.9)</b>
	Repetition, Parsing, Blending (34)	26	5	24	5	21	21	22	9	16	17	<b>16.6</b> <b>(7.7)</b>	22	10	13	11	14	20	14	96	26	9	0	<b>21.4</b> <b>(25.7)</b>
<b>Severity</b>		Moderate	Moderate	Mild	Severe	Moderate	Moderate	Mild	Severe	Mild	Severe	Mild (3)	Moderate	Moderate	Mild	Moderate	Mild	Mild	Mild	Mild	Moderate	Moderate	Mild (6)	Moderate (5)

Group A = individuals with AOS and concomitant aphasia;

Group P = individuals with aphasia without AOS

Three certified SLPs with 20 or more years of experience, proficient in the differential diagnosis of neurogenic communication disorders, served as expert raters for this study (secondary authors – KS, KY & DK). Expert raters observed audiovisual recordings of participant performance on subtests of the ABA and SRP to determine placement of participants into Group A or Group P. Consensus rating was employed; expert raters reached 100% agreement on each participant’s diagnosis.

Severity of impairment, or severity of spoken speech and/or language production abilities, was determined for each participant via consensus ratings by the three expert raters described above (Table 1). Expert raters watched audiovisual recordings of each participant completing subtest II from the ABA-2 and a 25-35 second speech sample from the SRP. Raters used their clinical knowledge to rate severity of impairment and were asked to use the following rating scale: 1 = mild, 2 = moderate, and 3 = severe. No other directions or criteria were provided. Ratings yielded three severity groups for Group A: mild ( $n = 3$ ), moderate ( $n = 4$ ), and severe ( $n = 3$ ) and two severity groups for Group P: mild ( $n = 6$ ) and moderate ( $n = 5$ ). Raters were in 100% agreement on each participant’s severity rating.

### ***Experimental Stimuli***

Repetition stimuli were created with the aim to induce errors in individuals with AOS (i.e., stimuli were motorically complex). Stimuli consisted of 30 two-, three-, and five-syllable real words containing consonant clusters and were separated into two 15-word lists (List A and B; see Appendix B). Each word list contained the same number of two-, three-, and five-syllable real words and were matched for psycholinguistic variables, including frequency of use (List A,  $M = 2.9$ ,  $SD = 3.8$ ; List B,  $M = 2.8$ ,  $SD = 3.9$ ), age of acquisition (List A,  $M = 453.8$ ,  $SD = 68.1$ ;

List B,  $M = 471.8$ ,  $SD = 86.3$ ), imageability (List A,  $M = 425.5$ ,  $SD = 78.6$ ; List B,  $M = 423.7$ ,  $SD = 100.1$ ), concreteness (List A,  $M = 374.0$ ,  $SD = 110.5$ ; List B,  $M = 383.1$ ,  $SD = 126.3$ ), familiarity (List A,  $M = 496.1$ ,  $SD = 31.0$ ; List B,  $M = 507.6$ ,  $SD = 56.8$ ), neighborhood density (List A,  $M = 1.47$ ,  $SD = 1.2$ ; List B,  $M = 1.5$ ,  $SD = 1.1$ ), and phonotactic probability (List A,  $M = .005$ ,  $SD = .003$ ; List B,  $M = .005$ ,  $SD = .002$ ). There were no significant differences between lists ( $p > .05$ ). Lists were also matched for part of speech, number of phonemes, and number of clusters. Lastly, articulatory features (voice, place, and manner), word onset, and overall syllable structure were similar across word lists.

A native English speaking adult male was selected to act as the “model” for stimulus elicitation. His productions of the 30 multisyllabic words were auditorily and visually recorded and consisted of a close up visual display of his articulators during production. Recordings were judged for accuracy and intelligibility by three members of the Aphasia Research Laboratory and were re-recorded as needed until 100% accuracy and intelligibility was reached.

### ***Experimental Procedures***

Participants were seated in a quiet room, either located in the UW Aphasia Research Laboratory or in their home. Stimuli were presented to participants via a computer monitor and external speakers (Alesis M1ACTIVE 520). For each trial, participants saw and heard a short video clip of a speaker’s mouth producing the target word one time. On the rare occurrence that a participant did not hear the word or were distracted during the initial presentation of the model, a repetition of the model was provided. A high quality head-mounted microphone (Audio-Technica ATM75) and video recorder (Canon VIXIA HF20) were utilized to capture participant productions for perceptual analysis.

Stimuli were elicited under two response conditions, blocked and random. The blocked condition consisted of 15 multisyllabic words (List A). Participants were asked to repeat each word, after a model, five times consecutively “as quickly and as clearly as possible.” Following participant production, the next target was then presented. The model for each word was presented only one time during the protocol. Target words were randomized within the word list for each participant. The random condition also consisted of 15 multisyllabic words (List B). Participants were asked to repeat the target word one time “as quickly and clearly as possible”. Following participant production, the next target was then presented. Each target was presented five times during the entire protocol. Target words were pseudo-randomized within the word list for each participant so that the same word was never presented twice in a row. Order of presentation condition was counter balanced across participants. Each condition was completed in a single session, which never exceeded 60 minutes. Feedback about production accuracy was not provided to the participant during the experimental task.

### ***Outcome Measures***

Four outcome measures: consistency of error location, variability of error type, error rate, and error type were determined by analysis of the five repetitions of the multisyllabic words (see Appendix C for a list of outcome measures and corresponding definitions and calculations). Consistency of error location was calculated to express the degree to which sound errors occur in the same target sound across five trials. It was defined as the number of instances in which the same sound segment was in error *three* times or more (Haley et al., 2013), and was calculated by dividing the number of sounds consistently in error (within a word type) by the number of total sounds in error ( $\# \text{ sounds consistently in error} / \# \text{ total sound errors}$ ; McNeil et al., 1995). The

computation provides a percentage, with high values indicating more consistent and low values indicating less consistent. Variability of error type was based on diacritic-level differences in the phonetic transcription. It refers to the degree to which sound errors differ from each other within the same location of a word, and was calculated by dividing the number of different errors that occurred within the same location of a target word across trials by the total number of errors produced within the same location of the target word across trials ( $\#$  of different errors /  $\#$  total errors within the same location; McNeil et al., 1995). This computation provides a percentage, with high values indicating more variability and low values indicating less variability. Error rate was expressed as the mean number of sound errors per word for each participant (Haley et al., 2013). Lastly, error type was expressed as the number of specified error types produced per participant. Errors were coded as follows: 1) distortions, 2) distorted substitutions, 3) distorted additions, 4) phonemic sequencing errors (i.e., perseverative, anticipatory, transposition errors), 5) non-sequential phonemic substitution errors, 6) omissions, and 7) additions. See Appendix D for a detailed description of each error type.

### ***Transcription and error coding***

All speech samples were analyzed perceptually utilizing narrow phonetic transcription via audiovisual recordings (following recommendations of Shriberg & Kent, 2003). The participants' first full realization of the target word, for each trial, was selected for analysis. Two groups of two transcribers, trained in narrow phonetic transcription (Shriberg & Kent, 2003), transcribed participant productions. In each group the first transcriber, TS1, was responsible for creating the original transcription and the second transcriber, TS2, was responsible for checking the original transcription (Du Bois, Cumming, Schuetze-Coburn, & Paolino, 1992). Consensus rules for

phonetic transcriptions were employed (Shriberg & Kent, 2003; Shirberg et al., 1984). Final transcriptions were coded for the seven predetermined error categories.

Productions that consisted of a non-response, semantic error, and/or a response that could not be reliably lined up with the target word (e.g., word had less than 30% of the target phonemes, but matched the number of target syllables, or word had less than 40% of the target phonemes and did not match the number of target syllables) were not included in the analysis.

### ***Reliability***

Reliability for narrow phonetic transcription and for error type coding was calculated using Cohen's Kappa. Fifteen percent of each participant's responses were randomly selected for re-analysis by trained transcribers. Inter-rater reliability for item-to-item agreement for narrow phonetic transcription was .74, and intra-rater was .80. These values indicate substantial agreement (Landis & Koch, 1977). Inter-rater reliability for item-to-item agreement for error type coding was .95, and intra-rater was .99. These values indicate almost perfect agreement (Landis & Koch, 1997). Lastly, intra-judge reliability for severity of impairment, performed on 10% of the data via percent agreement, was 100%.

### ***Statistical Analysis***

Research questions were primarily examined using nonparametric statistics, including Mann-Whitney U tests (research questions 1, 2c, and 3) and a Kruskal-Wallis H test (research question 2c). Research question 2b was addressed via visual analysis of scatter plots. The only parametric tests performed consisted of two two-way repeated analyses of variance (ANOVAs) for research

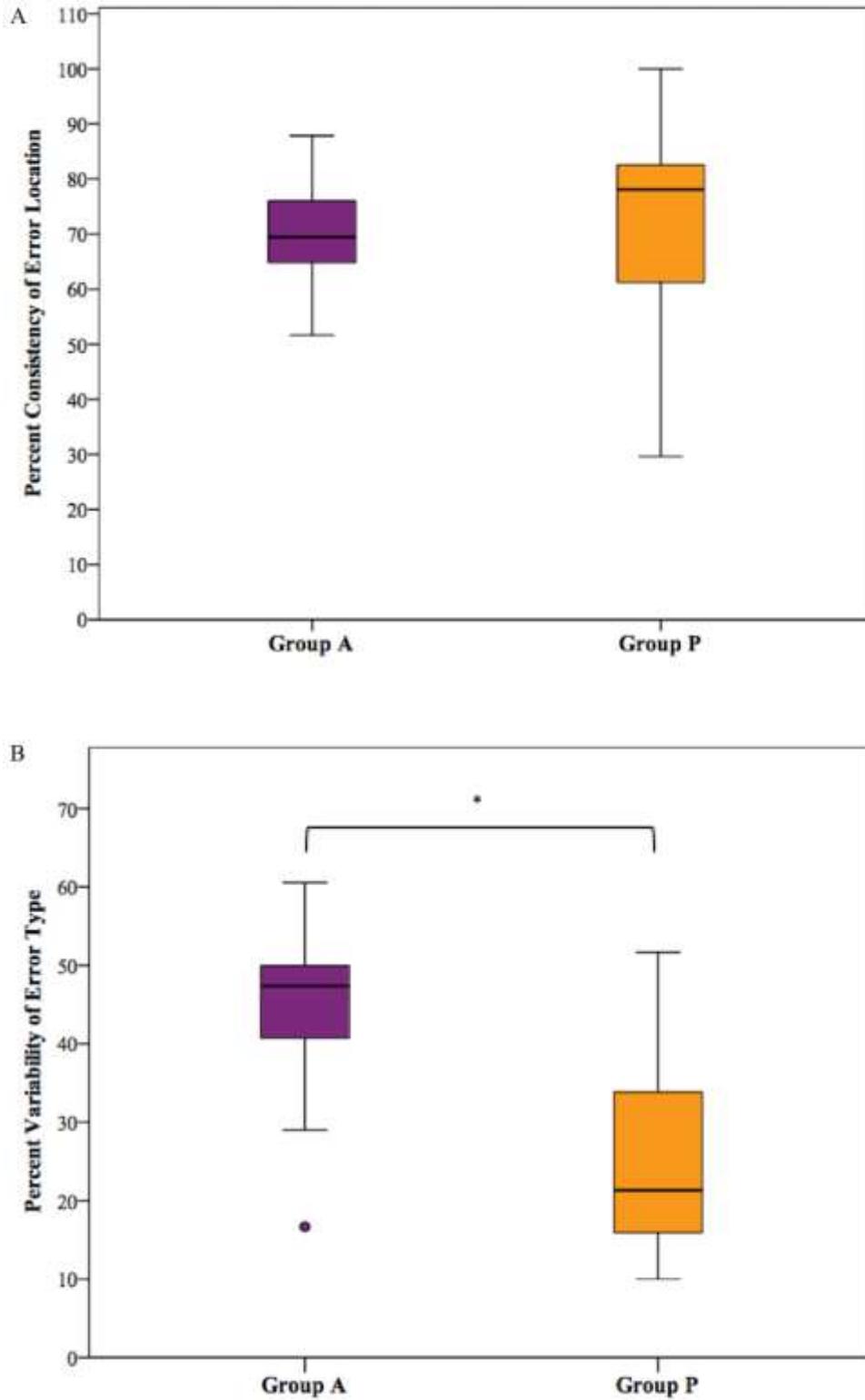
question 2a, in which all assumptions were met. Detailed descriptive statistics are provided in Appendix E, Tables a-d.

## Results

All 21 participants completed both blocked and random conditions for a total of 3,150 productions. Of these, 244 productions, including non-responses, semantic errors, and/or responses that could not be reliably lined up with the target word, were excluded from the study. Overall, 2,906 responses were included in the analyses reported below.

### *Research Question 1 – Error Consistency*

Mann-Whitney U tests were conducted to determine if there were differences in error consistency between Group A and Group P during the repetition of multisyllabic words in a blocked presentation condition (Figure 1a-b). Distributions of consistency of error location and variability of error type for Group A and Group P were not similar, as assessed by visual inspection. Values are mean rank unless otherwise stated. Consistency of error location for Group A (9.90) and Group P (12.00) was not statistically significantly different,  $U = 44.00$ ,  $z = -.775$ ,  $p = .468$ . Variability of error type for Group A (14.15) was statistically significantly higher than for Group P (8.14),  $U = 23.00$ ,  $z = -2.219$ ,  $p = .024$ .



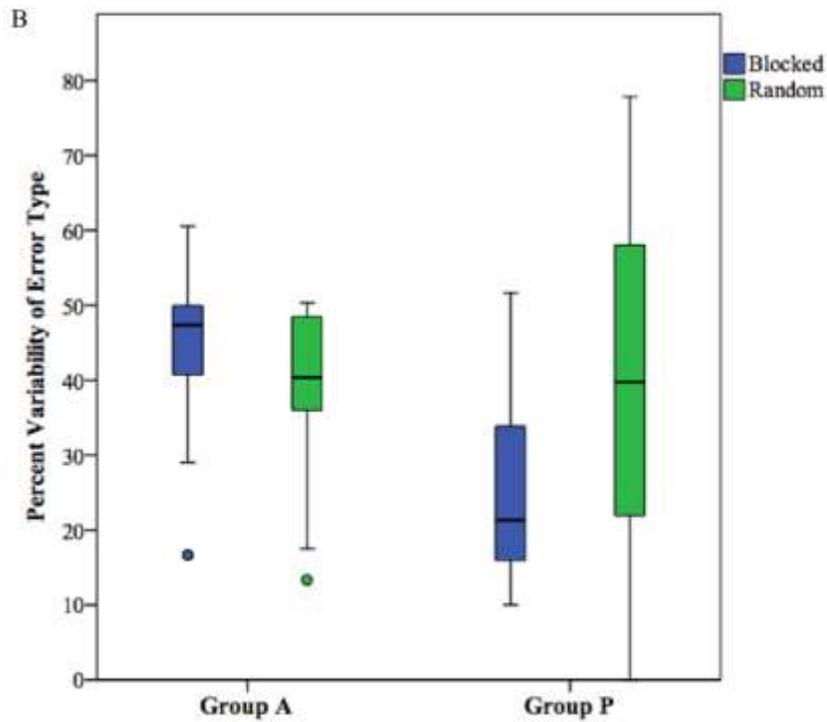
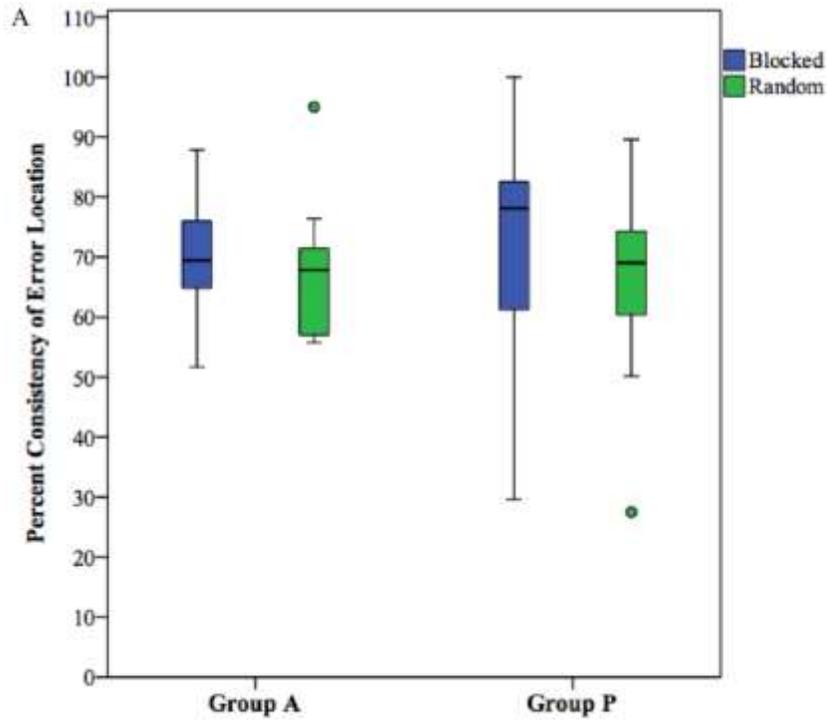
**Figure 1.** Group performance on (A) consistency of error location and (B) variability of error type in the blocked presentation condition. The primary (bold middle) horizontal bar denotes the median, and the extreme bars on each end denote the maximum and minimum values. Group A = individuals with acquired apraxia of speech (AOS) and concomitant aphasia; Group P = individuals with aphasia without acquired AOS; ○ = outlier, \* =  $p < .05$ .

## ***Research Question 2 – Influential Variables***

### *Research question 2a – Stimulus presentation condition.*

Two two-way repeated measures factorial ANOVAs were performed to examine the effects of group and stimulus presentation condition on error consistency during the repetition of multisyllabic words (Figure 2a-b). Analysis of the studentized residuals showed that there was normality, as assessed by the Shapiro-Wilk test of normality and no outliers (no studentized residuals greater than  $\pm 3$  standard deviations). No significant main effect was found for group,  $F(1, 9) = .040, p = .845$ , condition,  $F(1, 9) = .623, p = .450$ , and there was no interaction between the effect of group and condition on consistency of error location,  $F(1, 9) = .191, p = .672$ . No significant main effect was found for group,  $F(1, 9) = 2.588, p = .142$ , condition,  $F(1, 9) = .192, p = .671$ , and there was no interaction between the effect of group and condition on variability of error type,  $F(1, 9) = 2.313, p = .163$ .

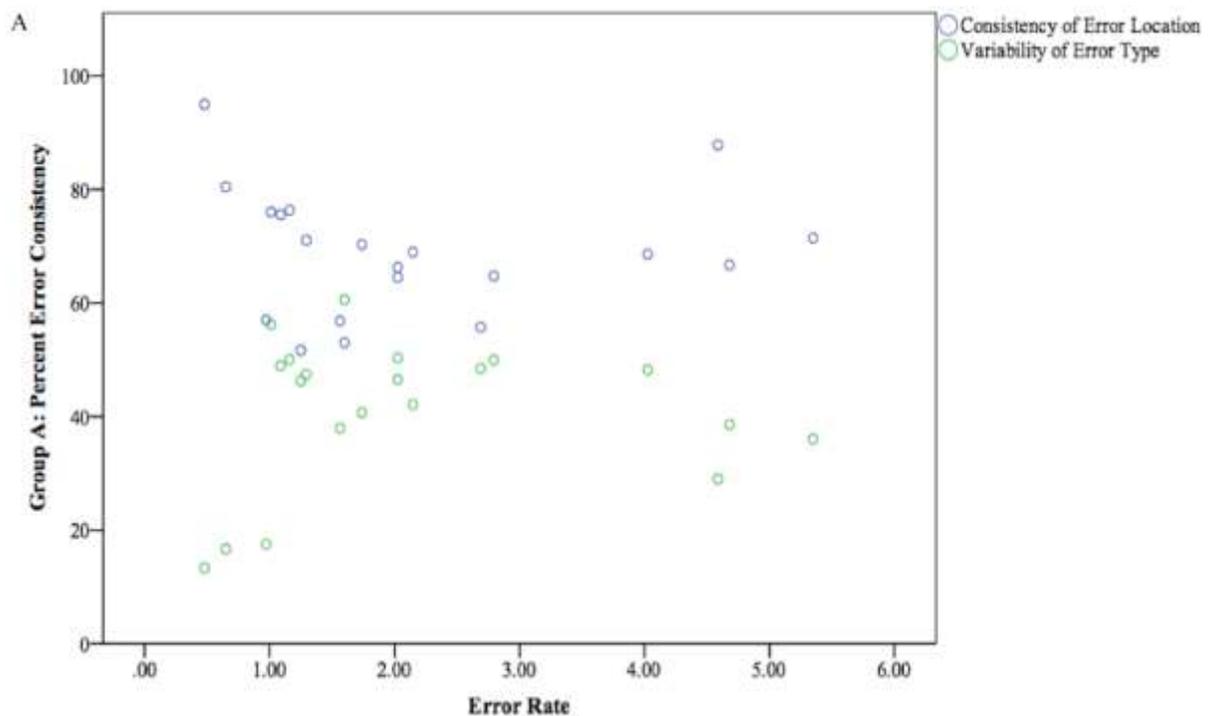
As a result of the findings reported above, stimulus presentation condition was collapsed across conditions for the analyses performed on research questions 2b and 2c.

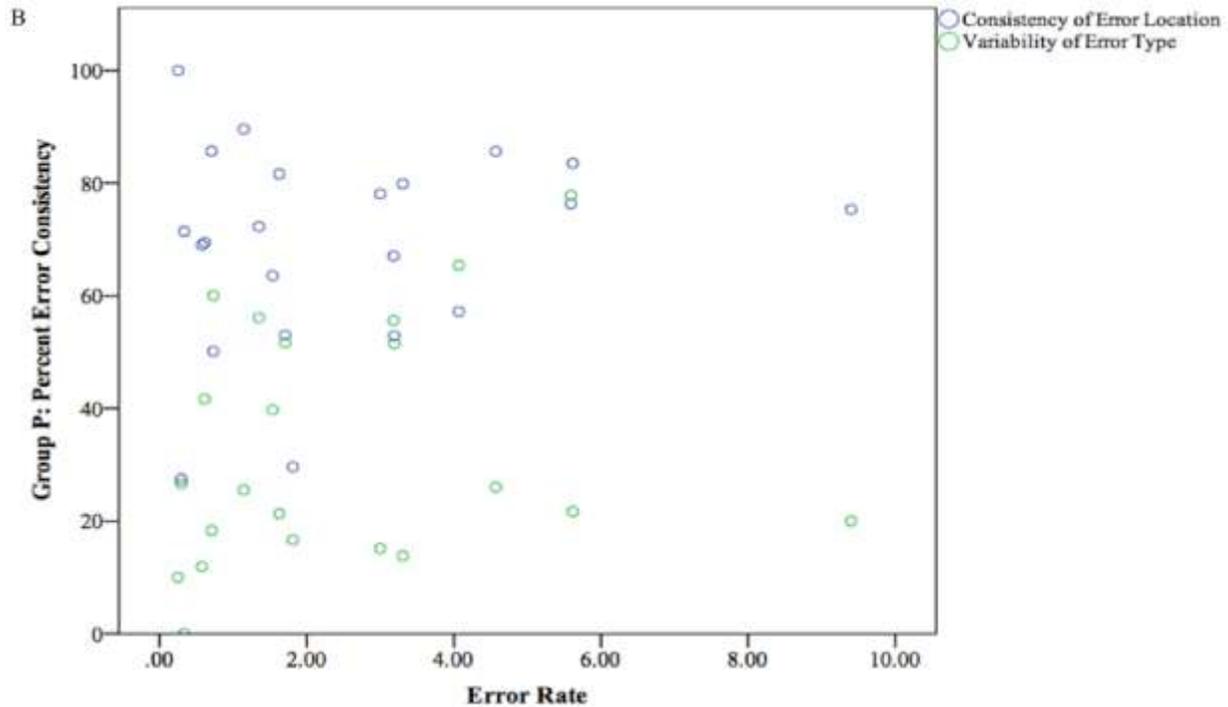


**Figure 2.** Effect of stimulus presentation condition on (A) consistency of error location and (B) variability of error type. The primary (bold middle) horizontal bar denotes the median, and the extreme bars on each end denote the maximum and minimum values; ○ = outlier.

*Research question 2b – Error rate.*

For Group A the overall error rate ranged from 0.48 to 5.35 incorrect sounds per word,  $M = 2.16$ ,  $SD = 1.43$ . For Group P the overall error rate ranged from 0.25 to 9.40 incorrect sounds per word,  $M = 2.48$ ,  $S = 2.29$ . Preliminary data showed the relationship for error rate and measures of error consistency for Group A and Group P to be non-linear, as assessed by visual inspection of a scatter plot (Figure 3a-b). For Group A, the data pattern is u-shaped, with a slope of zero, and a weak relationship between error rate and measures of error consistency. For Group P, the data pattern is mildly u-shaped, with a slope of zero, and a weak relationship between error rate and measures of error consistency.





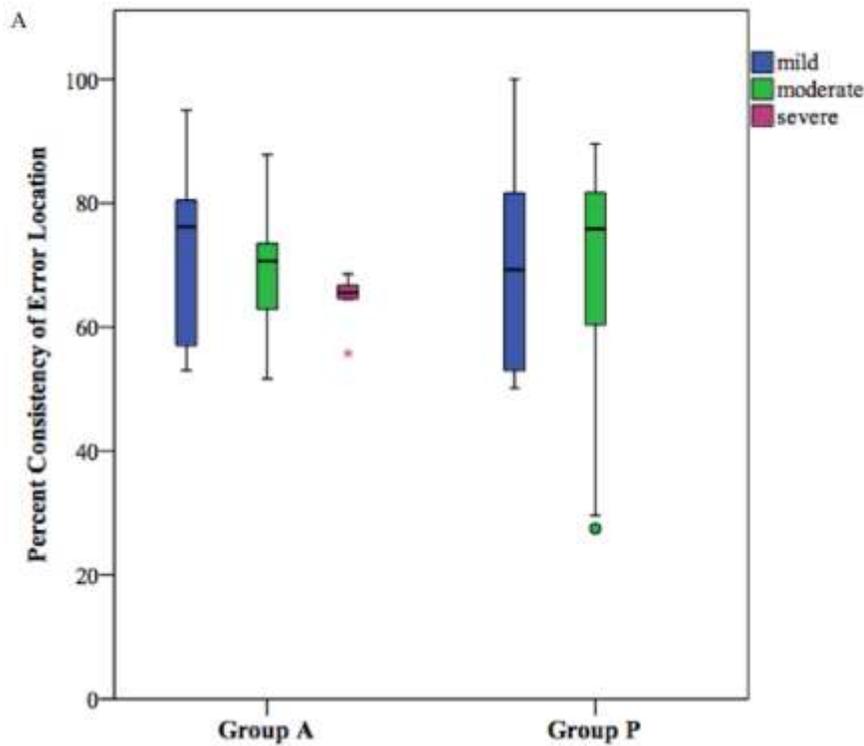
**Figure 3.** Influence of error rate on consistency of error location and variability of error type in (A) group A and (B) group P.

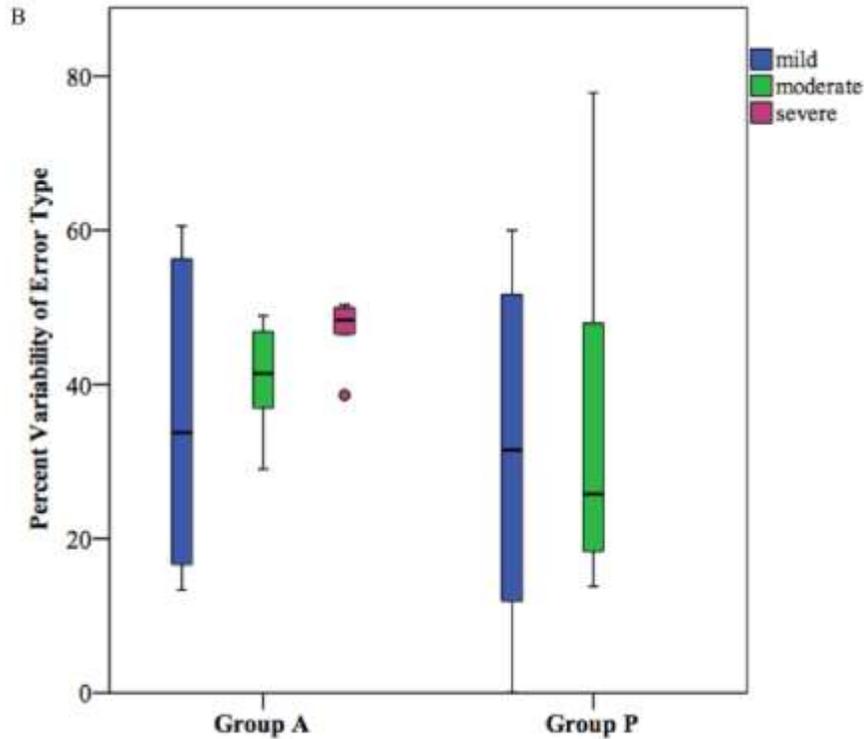
*Research question 2c – Severity of impairment.*

For Group A, a Kruskal-Wallis H test was run to determine if there was an influence of severity of impairment, “mild” (n=6), “moderate” (n=8), “severe” (n=6), on error consistency. Distributions of consistency of error location were similar for the mild and moderate severity group, but not the severe group, as assessed by visual inspection of a boxplot. Distributions of variability of error type were not similar for the three severity groups (mild, moderate, severe). The distributions of outcomes for consistency of error location and variability of error type were not statistically significantly different between groups,  $\chi^2(2) = 3.065$ ,  $p = .216$  and  $\chi^2(2) = 2.027$ ,  $p = .363$ , respectively (Figure 4a-b).

For Group P Mann-Whitney U tests were performed, with a bonferroni correction for multiple comparisons, to determine if there were within group differences in error consistency

between severity groups, “mild” (n=12) and “moderate” (n = 10). Statistical significance was accepted as the  $p < .0125$  level. Distributions of consistency of error location and variability of error type for severity categories were not similar, as assessed by visual inspection. Values are mean ranks unless otherwise stated. Consistency of error location for the mild (12.58) and moderate (10.20) severity groups were not statistically significantly different,  $U = 47.000$ ,  $z = -.857$ ,  $p = .418$ . Variability of error type for mild (9.08) and moderate (14.40) severity groups were not statistically significantly different,  $U = 31.00$ ,  $z = -1.912$ ,  $p = .59$  (Figure 4a-b).



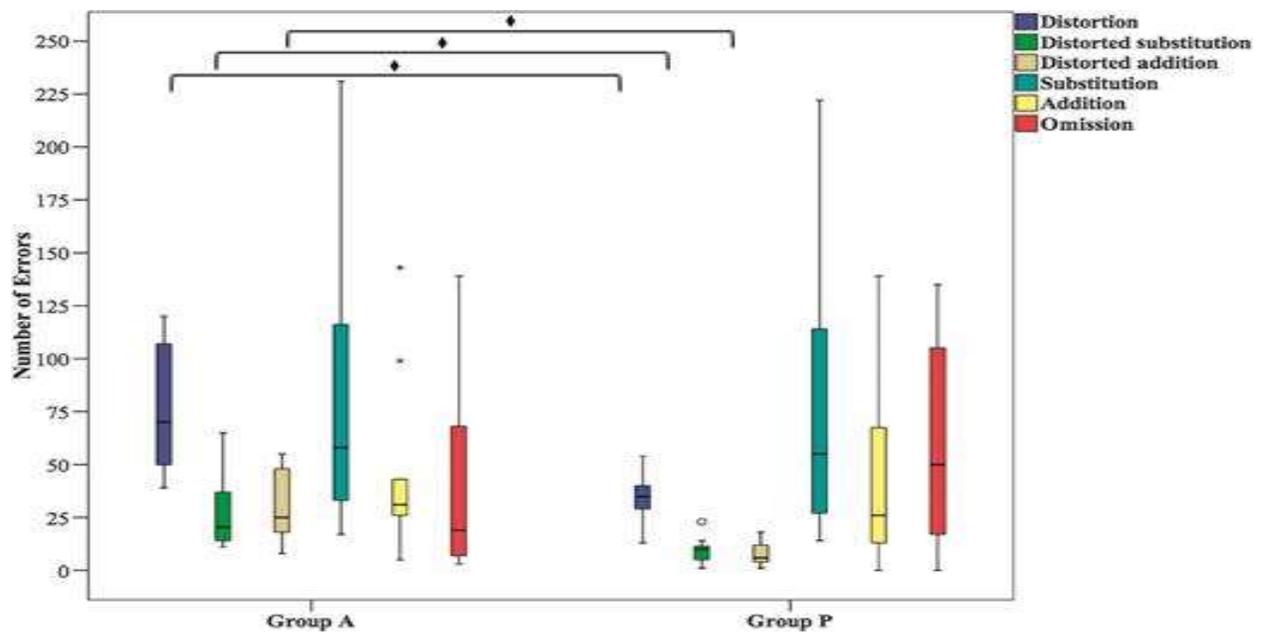


**Figure 4.** Influence of severity of impairment on (A) consistency of error location and (B) variability of error type; ○ = outlier, \* = extreme outlier.

### ***Research Question 3 – Error Type***

Mann-Whitney U tests were conducted, with a bonferroni correction for multiple comparisons, to determine if there were differences in the types of errors produced between Group A and Group P in the blocked condition (Figure 5). Statistical significance was accepted at the  $p < .01$  level. Distributions of error type, for all error type categories, for Group A and Group P were not similar, as assessed by visual inspection. Values are mean ranks unless otherwise stated. Due to the small occurrence of sequential substitution errors in Group A ( $M = 9.20$ ,  $SD = 4.80$ ) and Group P ( $M = 6.00$ ,  $SD = 4.31$ ), along with similarities in group performance, the sequential substitution category was collapsed with the non-sequential substitution category, now referred to as “substitutions”.

The occurrence of distortions (Group A = 15.95, Group P = 6.50;  $U = 5.50$ ,  $z = -3.487$ ,  $p = .000$ ), distorted substitutions (Group A = 15.50, Group P = 6.91;  $U = 10.00$ ,  $z = -3.178$ ,  $p = .001$ ), and distorted additions (Group A = 15.65, Group P = 6.77;  $U = 8.50$ ,  $z = -3.284$ ,  $p = .000$ ) were significantly higher for Group A than Group P. Groups did not perform significantly different in the occurrence of substitutions (Group A = 11.30, Group P = 10.73;  $U = 52.00$ ,  $z = -.211$ ,  $p = .833$ ), additions (Group A = 12.00, Group P = 10.09;  $U = 45.00$ ,  $z = -.705$ ,  $p = .481$ ), and omissions (Group A = 9.55, Group P = 12.32;  $U = 40.50$ ,  $z = -1.012$ ,  $p = .307$ ).



**Figure 5.** Error types produced in the blocked condition; ° = outlier, \* = extreme outlier, ♦ =  $p < .01$ .

## Discussion

The primary aim of this study was to examine reliability of error location and type in a group of individuals with AOS and concomitant aphasia compared to individuals with aphasia without AOS. To that end, the consistency of error location and variability of error type was examined in 21 participants during repetition of 30 multisyllabic words across five trials. The influence of stimulus presentation condition, error rate, and severity of impairment on group performance was

also explored, as well as group differences in the types of errors produced. Overall, results do not support current diagnostic guidelines identifying relatively consistent errors as a primary characteristic of AOS, nor do they support the original belief that errors in AOS are more variable than PP. Furthermore, findings do not support relative error consistency, or inconsistency, as effective criteria in the differential diagnosis of these two clinical populations.

### ***Research Question 1 – Error Consistency***

#### *Consistency of Error Location*

Individuals with AOS and concomitant aphasia performed similarly to individuals with aphasia without AOS on the consistency of error location. These findings differ from those of McNeil et al. (1995), but are consistent with those reported by Haley and colleagues (2013). Similar performance between groups suggests that the presence of AOS in Group A did not influence consistency of error location. In addition, the complex stimuli employed in this investigation also may have led to more reliable error production in both groups. Stimuli were created with the aim to induce errors in Group A by increasing the motoric complexity of the target; consequently, stimuli may have also been linguistically challenging for Group P. Indeed, past studies suggest that word length (e.g., Kohn, 1989; Nickels, 1995; Shallice, Rumiati, & Zadini, 2000) and number of phonemes influence phonological errors in individuals with aphasia (Nickels & Howard, 2004). In the context of Dell's model, the more phonemes that require activation, the more the processing load, the greater the time course required for activation spread and opportunities for interference (Dell, 1988).

### *Variability of error type*

Group A demonstrated higher variability of error type compared to Group P in the blocked presentation condition. This finding most likely reflects the co-occurrence of AOS and aphasia in Group A, in that the errors produced result from damage to two different mechanisms, linguistic processing and speech motor planning/programming, and numerous potential sources of disruption (e.g., disordered spreading activation, variability of feedforward control; Dell, 1986, 1988; Guenther, 2006), allowing for a greater range of error types compared to individuals with aphasia without AOS. These findings, however, are not consistent with participant performance in the random presentation condition (see research question 2a).

Taken together, these results suggest that relatively consistent errors should not be considered a primary or differentiating characteristic of AOS in individuals with AOS and concomitant aphasia.

### ***Research Question 2 – Influential Variables***

#### *Research question 2a – Stimulus presentation condition*

Stimulus presentation condition did not influence performance on measures of error location and type for either group. This finding was somewhat unexpected given the theoretical support for differences in the retrieval of stimuli in blocked vs. random presentation conditions (Dell, 1986; Knock et al., 2000; Maas, et al., 2008; Schmidt & Lee, 2005). Results of Mauszycki and colleagues (2010a, 2010b, 2012), however, are consistent with the findings reported here.

Notably, examination of individual data for Group P shows patterns of individual differences that were not observed in Group A (see Appendix F for individual data). Specifically, for Group P half or more participants demonstrated more reliable errors in the blocked condition compared

to the random condition, for both error location and type. Others demonstrated minimal or no effect of presentation condition, and even fewer participants demonstrated more reliable performance in the random condition. In contrast, individuals in Group A were more similar in performance across presentation conditions (with the exception of Pt 9 on variability of error type). The individual variability observed in Group P may reflect the heterogeneity of aphasia included in this group. For some individuals the activation of new phonological representations in the random condition may provide greater opportunity for interference, therefore, leading to more variable productions compared to the blocked condition. For others, specifically those with relatively poor maintenance of activation (i.e., impaired verbal working memory), the blocked condition may lead to more variable productions compared to the random condition, as activation quickly decays across trials (Martin & Gupta, 2004). A closer look at individual performance would be beneficial in understanding underlying mechanisms that contribute to the heterogeneity within this population.

#### *Research question 2b – Error rate*

In the present study scatter plots revealed a weak association between error rate and measures of error consistency for both groups. In particular, the data pattern for Group A suggests that individuals with high and low error rates have higher consistency of error location and lower variability of error type, while individuals with more intermediate error rates have lower consistency of error location and higher variability of error type. Interestingly, Haley and colleagues (2013) report a relationship between more frequent errors and higher consistency of error location and also higher variability of error type. The results of Haley et al. (2013) reflect the performance of 32 participants with aphasia and/or AOS (and possibly concomitant mild

dysarthria), separated into groups based on error frequency (e.g., minimal errors vs more frequent errors). In contrast, the current study examined 21 participants, separated into groups based on diagnosis (Group A and Group P). Therefore, methodological differences between these two studies may be responsible for differences in the overall findings.

### *Research question 2c – Severity of impairment*

Similar to the findings reported above, severity of impairment did not appear to influence measures of error consistency in either group. It is worth mentioning, however, that the mild and moderate severity groups in Group A demonstrate a greater range of error consistency compared to the severe severity group (Figure 4a-b). Unfortunately, severity groups were not balanced across Group A and Group P. The range of severity perceived in Group A vs Group P (e.g., mild and moderate only) may reflect the combined effects of motor and linguistic deficits. Feedback from expert raters indicate that for Group A mild ratings were attributed to motor characteristics alone (e.g., slow speech rate, abnormal prosody, sound distortions), whereas more severe ratings were attributed to a combination of linguistic (e.g., empty or non-specific speech, agrammatism, anomia) and more significant motor characteristics (e.g., extreme slow rate, segmented speech, lengthened segments, impaired prosody, significantly impaired speech output).

### *Research Question 3 – Error Type*

As anticipated and guaranteed by the selection criteria, Group A produced significantly more phonetic errors compared to Group P. This finding is consistent with previous literature that identifies AOS as a motor speech disorder characterized by distortions, distorted substitutions

(e.g., Mauszycki et al., 2010a, 2010b, 2012; Odell et al., 1990), and distorted additions (Duffy et al., 2015).

Results demonstrate that individuals in Group P also produced phonetic errors. There are multiple reasons for this finding. First, two participants in Group P demonstrated excessive effort (e.g., straining and tension) as a result of anxiety and frustration during the experimental task. Data from these two participants contributed significantly to the phonetic errors observed in Group P. Secondly, a handful of studies examining speech kinematics and speech timing in AOS and aphasia suggest that there may be a phonetic-motoric component contributing to speech patterns in individuals with conduction aphasia (e.g., McNeil & Adams, 1991; McNeil, Liss, Tseng, & Kent, 1990), which may have contributed to the observed speech errors. Thirdly, neurologically healthy adults occasionally produce phonetic errors, most often during difficult speech tasks, such as tongue twister paradigms (Goldrick & Blumstein, 2006), speeded repetition of word pairs (Goldstein, Pouplier, Chen, Saltzman, & Byrd, 2007), and SLIP (spoonerisms of laboratory induced predisposition; Pouplier, 2007). Factors contributing to phonetic breakdown in individuals without motor speech impairments may include fast or rushed speech rate and/or phrases with competing or similar sounds. In the current study, participants were asked to repeat complex multisyllabic words five times in a row “as quickly and as clearly as possible.” This elicitation method, in combination with complex stimuli, may have facilitated the production of phonetic errors in individuals with brain damage and potential resource allocation deficits.

Finally, groups did not differ in the number of substitution, addition, and omission errors. This finding can likely be attributed to the occurrence of aphasia with PP in each group. Furthermore, non-sequential substitution, omission, and addition errors have been observed in individuals with isolated AOS (e.g., Itoh, Sasanuma, & Ushijima, 1979; Odell et al., 1990).

## **Conclusion**

Accurate and reliable diagnosis of AOS continues to be challenging for clinicians. Confusion surrounding the primary characteristics of the disorder and continued use of outdated diagnostic criteria likely contribute to this challenge. Therefore, studies improving upon current diagnostic criteria and procedures are of great value. The results of this investigation provide further evidence that relative error consistency is not a valid metric with which to diagnose AOS or differentiate individuals with AOS and aphasia from aphasia without AOS. The external validity of the present study is strengthened by the use of a participant sample that demonstrates a range of impairment and is representative of the clinical population served by SLPs.

### ***Clinical Implications***

As a result of these findings, it is recommended that the characteristic of relative error consistency, or relative error inconsistency, not be considered a primary characteristic of AOS. Furthermore, findings support the continued use of phonetic errors in the characterization and diagnosis of AOS. Importantly, distortion errors occurred in individuals with aphasia without AOS, and these errors are also considered a primary characteristic of dysarthria. Thus, the observation of distortion errors alone does not warrant a diagnosis of AOS, rather it is the combination or clustering of multiple speech characteristics (e.g., distortions, slow rate, abnormal prosody, segment segregation, etc.).

### ***Study Limitations and Future Directions***

Sample size is often an issue in investigations focused on rare or special populations. The current investigation consisted of 21 participants separated into two diagnostic groups, Group A -

ten individuals with AOS and concomitant aphasia, and Group P - 11 individuals with aphasia without AOS. Diagnostic groups were divided into even smaller subgroups when examining the influence of severity of impairment on error consistency. A larger sample size would have provided more power and permitted more advanced statistical procedures (e.g., multivariate linear or logistic regression).

Future work aims to expand on that of Odell and colleagues (1990; 1991) with a more detailed investigation of error type, extending beyond error categories to examine error type patterns, such as prominent types of distortion and substitution errors (e.g., voicing errors), as well as within-word variables that influence breakdown, such as word and syllable location (e.g., onset vs. coda, singletons vs. clusters). Lastly, examination of learning across repeated trials, without feedback, is also of interest, as there is both theoretical (Guenther, 2006; Schmidt & Lee, 2005) and patient (Buchwald & Miozzo, 2012) data to support improved speech production in individuals with AOS. This information may contribute to the diagnostic process and inform treatment development.

### **Acknowledgments**

This research was supported by National Institute on Deafness and Other Communication Disorders Grant F31DC 013947, awarded to Lauren Bislick. The authors acknowledge Melissa Helen, Christi Quilligan, Lisa Bjorback, Katrina Ross, Laura Cain, Josie Stump and all participants and their families for their contributions to this study.

## References

- Aichert, I. & Ziegler, W. (2008). Learning a syllable from its parts: Cross-syllabic generalisation effects in patients with apraxia of speech. *Aphasiology*, 22(11), 1216–1229.  
doi:10.1080/02687030701820303
- Brookshire, C.E., Conway, T., Hunting Pompon, R., Oelke, M., & Kendall, D. (2014). Effects of intensive phonomotor treatment on reading in eight individuals with aphasia and phonological alexia. *American Journal of Speech-Language Pathology*, 23(10): S300-S311. doi:  
10.1044/2014\_AJSLP-13-0083
- 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), 1573–86. doi:10.1044/1092-4388(2012/11-0200)
- Dabul, B. (2000). *Apraxia battery for adults (second ed.)*. Austin, Tx: Pro-Ed.
- Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, 93(3), 283–321. doi:10.1037/0033-295X.93.3.283
- Dell, G. S. (1988). The retrieval of phonological forms in production: Tests of predictions from a connectionist model. *Journal of Memory and Language*, 27(2), 124–142. doi:10.1016/0749-596X(88)90070-8
- Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review*, 104, 801–939.
- Du Bois, J. W., Cumming, S., Schuetze-Coburn, S., & Paolino, D. (1992). Discourse transcription. *Santa Barbara Papers in Linguistics*, 4. <http://www.linguistics.ucsb.edu/research/santa-barbara-papers>
- Duffy, J. R. (2013). *Motor speech disorders: Substrates, differential diagnosis and management*, 3rd Edition. St. Louis, Mo. Elsevier Mosby.

- Duffy, J. R., Strand, E. A., Clark, H., Machulda, M., Whitwell, J., & Keith, J. A. (2015). Primary progressive apraxia of speech: Clinical features and acoustic and neurologic correlates. *American Journal of Speech-Language Pathology, 24*, 88–100. doi:10.1044/2015\_AJSLP-14-0174
- Goldrick, M. & Blumstein, S. E. (2006). Cascading activation from phonological planning to articulatory processes: Evidence from tongue twisters. *Language and Cognitive Processes, 21*, 649-683. doi:10.1080/01690960500181332
- Goldstein, L., Pouplier, M., Chen, L., Saltzman, E., & Byrd, D. (2007). Dynamic action units slip in speech production errors. *Cognition, 103*(3), 386-412. doi:10.1016/j.cognition.2006.05.010
- Guenther, F. H. (2006). Cortical interactions underlying the production of speech sounds. *Journal of Communication Disorders, 39*(5), 350-365. doi: 10.1016/j.jcomdis.2006.06.013
- Guenther, F. H., Ghosh, S. S., & Tourville, J. A. (2006). Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain and Language, 96*(3), 280-301. doi: 10.1016/j.bandl.2005.06.001
- Haley, K. L., Bays, G. L., & Ohde, R. N. (2001). Phonetic properties of aphasic-apraxic speech: A modified narrow transcription analysis. *Aphasiology, 15* (12), 1125-1142. doi:10.1080/02687040143000537
- 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*, 891-905. doi:10.1044/1092-4388(2012/12-0161)
- Itoh, M., Sasanuma, S., & Ushijima, T. (1979). Velar movements during speech in a patient with apraxia of speech. *Brain and Language, 7*, 227–239. doi:10.1016/0093-934X(79)90019-1
- Jacks, A. (2008). Bite block vowel production in apraxia of speech. *Journal of Speech, Language, and Hearing Research, 51*, 898-913. doi:10.1044/1092-4388(2008/066)

- Johns, D. F., & Darley, F. L. (1970). Phonemic variability in apraxia of speech. *Journal of Speech and Hearing Research, 13*(3), 556–83. doi:10.1044/jshr.1303.556
- Kendall, D. L., del Toro, C. M., Nadeau, S., Johnson, J., Rosenbek, J., & Velozo, C. (2010). The development of a standardized assessment of phonology in aphasia. Presented at the Clinical Aphasiology Conference, Isle of Palms, South Carolina.
- Kent, R. & Rosenbek, J. (1983) Acoustic Patterns of apraxia of speech. *Journal of Speech and Hearing Research, 26*, 231-249. doi:10.1044/jshr.2602.231
- Knock, T., Ballard, K., Robin, D., & Schmidt, R. (2000). Influence of order of stimulus presentation on speech motor learning: A principled approach to treatment for apraxia of speech. *Aphasiology, 14*(5-6), 653–668. doi:10.1080/026870300401379
- Kohn, S. E. (1989). The nature of the phonemic string deficit in conduction aphasia. *Aphasiology, 3*, 209-239. 10.1080/02687038908248992
- Landis, J. R. & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159-174. doi: 10.2307/2529310
- LaPointe, L. L. & Johns, D. F. (1975). Some phonemic characteristics in apraxia of speech. *Journal of Communication Disorders, 8*(3), 259–69. doi:10.1016/0021-9924(75)90018-0
- Maas, E., Mailend, M.-L., & Guenther, F.H. (2015). Feedforward and feedback control in apraxia of speech: Effects of noise masking on vowel production. *Journal of Speech, Language, and Hearing Research, 53*, 185-200. doi:10.1044/2014\_JSLHR-S-13-0300
- Maas, E., Austerman Hula, S., Robin, D. A., Freedman, S. E., Ballard, K. J., & Schmidt, R. A. (2008). Principles of motor learning in treatment of motor speech disorders. *American Speech-Language-Hearing Association, 17*(3), 277–98. doi:10.1044/1058-0360(2008/025)
- Mackay, D. G., & James, L. E. (2004). Sequencing, speech production, and selective effects of aging on phonological and morphological speech errors. *Psychology and Aging, 19*(1), 93–107. doi:10.1037/0882-7974.19.1.93

- Marquardt, T. P., Schneider, H. G., & Jacks, A. (2010). Error prediction in acquired apraxia of speech. *Journal of Medical Speech-Language Pathology*, *18*(4), 83–88.
- Martin, N. & Gupta, P. (2004). Exploring the relationship between word processing and verbal short-term memory: Evidence from associations and dissociations. *Cognitive Neuropsychology*, *21*, 213-228. doi: 10.1080/02643290342000447
- Mauszycki, S. C., Dromey, C., & Wambaugh, J. L. (2007). Variability in apraxia of speech: A perceptual, acoustic, and kinematic analysis of stop consonants. *Journal of Medical Speech-Language Pathology*, *15*, 223–242.
- Mauszycki, S. C., & Wambaugh, J. L. (2006). Perceptual analysis of consonant production in multisyllabic words in apraxia of speech: A comparison across repeated sampling times. *Journal of Medical Speech-Language Pathology*, *14*, 263–268.
- Mauszycki, S. C., Wambaugh, J. L., & Cameron, R. M. (2010a). Variability in apraxia of speech: Perceptual analysis of monosyllabic word productions across repeated sampling times. *Aphasiology*, *24*(6-8), 838–855. doi:10.1080/02687030903438516
- Mauszycki, S. C., Wambaugh, J. L., & Cameron, R. M. (2010b). Apraxia of speech: Perceptual analysis of bisyllabic word productions across repeated sampling occasions. *Journal of Medical Speech-Language Pathology*, *18*(4), 89–98.
- Mauszycki, S. C., Wambaugh, J. L., & Cameron, R. M. (2012). Apraxia of speech: perceptual analysis of trisyllabic word productions across repeated sampling occasions. *American Journal of Speech-Language Pathology*, *21*(2), 28–37. doi:10.1044/1058-0360(2011/11-0094)
- McNeil, M. R., Adams, S. (1991). A comparison of speech kinematics among apraxic, conduction aphasic, ataxic dysarthric, and normal geriatric speakers. In T.E. Prescott (Ed.), *Clinical aphasiology*, *19*, 279-294. Austin, TX:PRO-ED.

- McNeil, M. R., Liss, J., Tseng, T-H., & Kent, R. D. (1990). Speech timing in apraxia and conduction aphasia: A phonetic-motoric disorders. *Brain and Language*, 38, 135-158.
- 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 and differential diagnosis. In M. R. McNeil (Ed.), *Clinical management of sensorimotor speech disorders* (pp. 249-268). New York: Thieme.
- McNeil, M. R., Sung, J. E., Yang, D., Pratt, S. R., Fossett, T. R. D., Doyle, P. J., & Pavelko, S. (2007). Comparing connected language elicitation procedures in persons with aphasia: Concurrent validation of the Story Retell Procedure. *Aphasiology*, 21(6-8), 775-790.  
doi:10.1080/02687030701189980
- Nickels, L. A. (1995). Getting it right? Using aphasicnaming errors to evaluate theoretical models of spoken word production. *Language and Cognitive Processes*, 10, 13–45.  
doi:10.1080/01690969508407086
- Nickels, L., & Howard, D. (2004). Dissociating effects of number of phonemes, number of syllables, and syllabic complexity on word production in aphasia: It's the number of phonemes that counts. *Cognitive Neuropsychology*, 21(1), 57–78. doi:10.1080/02643290342000122
- Odell, K., McNeil, M. R., Rosenbek, J. C., & Hunter, L. (1990). Perceptual characteristics of consonant production by apraxic speakers. *Journal of Speech and Hearing Disorders*, 55(2), 345–59. doi:10.1044/jshd.5502.345
- 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. doi:10.1044/jshr.3401.67

- Pouplier, M. (2007). Tongue kinematics during utterances elicited with the SLIP technique. *Language and Speech*, 50, 311-341. doi: 10.1177/00238309070500030201
- Raven, J., Raven, J. C., & Court, J. H. (1998). *Manual for Raven's progressive matrices and vocabulary scales*. Oxford: Oxford Psychologists.
- Schmidt, R. A. & Lee, T. D. (2005). *Motor control and learning: a behavioral emphasis*. Champaign, IL: Human Kinetics.
- Shallice, T., Rumiati, R. I., & Zadini, A. (2000). The selective impairment of the phonological output buffer. *Cognitive Neuropsychology*, 17, 517–546. doi:10.1080/02643290050110638
- Shriberg, L. D., & Kent, R. D. (2003). *Clinical phonetics*. Boston, MA: Allyn and Bacon.
- Shirberg, L. D., Kwiatkowski, K., & Hofmann, J. (1984). A procedure for phonetic transcription by consensus. *Journal of Speech and Hearing Research*, 27, 456-465.
- Shuster, L. I., & Wambaugh, J. L. (2008). Token-to-token variability in adult apraxia of speech: A perceptual analysis. *Aphasiology*, 22(6), 655–669. doi:10.1080/02687030701632161
- Square, P. A., Darley, F., & Sommers, R. (1982). An analysis of productive errors made by pure apraxic speakers with differing loci of lesions. In R. Brookshire (Ed.), *Clinical aphasiology conference proceedings* (pp.245-250). Minneapolis, MN: BRK.
- Staiger, A., Finger-Berg, W., Aichert, I., & Ziegler, W. (2012). Error variability in apraxia of speech: a matter of controversy. *Journal of Speech, Language, and Hearing Research*, 55(5), 1544–61. doi:10.1044/1092-4388(2012/11-0319)
- Strand, E. A., & McNeil, M. R. (1996). Effects of length and linguistic complexity on temporal acoustic measures in apraxia of speech. *Journal of Speech and Hearing Research*, 39(5), 1018–33. doi:10.1044/jshr.3905.1018
- Swinburn, K., Porter, G., & Howard, D. (2004). *The comprehensive aphasia test*. Hove, East Sussex: Psychology Press.

- van der Merwe, A. (2009). A theoretical framework for the characterization of pathological speech sensorimotor control. In M. R. McNeil (Ed.), *Clinical management of sensorimotor speech disorders* (2nd ed.). New York: Thieme.
- 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), xv–xxxiii.
- Wambaugh, J. L., Nessler, C., Bennett, J., & Mauszycki, S. C. (2004). Variability in apraxia of speech: A perceptual and VOT analysis of stop consonants. *Journal of Medical Speech-Language Pathology*, *12*, 221–228.
- Yorkston, K., Beukelman, D.R., Strand, E. & Hakel, M. (2010). *Management of Motor Speech Disorders in Children and Adults*, Pro-Ed, Austin, Tex, USA
- Ziegler, W., Aichert, I., & Staiger, A. (2012). Apraxia of speech: Concepts and controversies. *Journal of Speech, Language, and Hearing Research*, *55*(5), 1485–501. doi:10.1044/1092-4388(2012/12-0128)

## Appendix A

### Participant Demographics for Groups A and P

Group and Participant	Age (years)	Handedness	Education (years)	MPO	Gender	Location of Stroke
Group A						
A1	61	R	18	172	F	Left M2 segment region infarct involving the lenticulostriate system, and also insular cortex.
A2	58	R	19	100	M	Left MCA infarct, involving frontal, temporal, and parietal lobes.
A3	68	R	23	50	M	Left MCA, involving the basal ganglia, adjacent insular cortex, left corona radiata, and left frontal lobe.
A4	67	R	18	230	M	Large left MCA.
A5	51	R	14	40	F	Two strokes: Left MCA aneurysmal subarachnoid hemorrhage, including left circular sulcus, sylvian fissure, and operculum extending deep to the cavernous sinus region; and left MCA occlusive event with infarction of temporal, frontal, and parietal lobes.
A6	66	R	16	136	M	Encephalomalacia evident left frontal distribution consistent with sequelae of remote infarct.
A7	60	R	17	42	F	Large left M1 infarction, with extensive frontal and temporal lobe involvement, including the superior temporal gyrus, and extensive involvement of the insula.
A8	45	R	16	10	F	Moderate left MCA infarct, including the posterior inferior lateral left parietal lobe with mild involvement of the posterior insula and left caudate head.
A9	71	R	16	22	F	Left M2 region infarct with loss of left insular ribbon, loss of gray white differentiation in the left frontal operculum and left parietotemporal lobes.
A10	71	R	13	21	F	Left MCA infarct involving the temporal lobe, posterior frontal lobe, and anterior parietal lobe. Minimal periventricular deep white matter T2 signal alteration, extending to involved insular cortex and posterior temporoparietal cortex and inferior frontal cortex.
<i>Avg (SD)</i>	61.8 (8.2)	10 Right	17 (2.7)	82.3 (70.8)	4 M, 6 F	
Group P						
P1	63	R	16	52	M	Left territory infarct involving the superior temporal gyrus, dorsal insula, and frontal operculum, extending deep rostrally to anterior periventricular white matter.
P2	70	R	12	68	F	Two strokes: Left MCA territory infarcts, including the temporoparietal junction, coronal radiata, and subinsular region.
P3	57	R	12	57	F	Large left MCA territory infarcts, including much of the temporal and frontal lobes.
P4	49	R	16	61	M	Small left cortical infarct, including lateral frontal

Group and Participant	Age (years)	Handedness	Education (years)	MPO	Gender	Location of Stroke
P5	59	R	18	110	F	lobe. Left MCA territory infarct, including frontal, temporal, and parietal lobes, prior to aneurysm clipping.
P6	57	R	16	65	M	Left hemorrhagic, including basal ganglia, post frontal, temporal, and parietal lobes.
P7	66	R	14	125	F	Extensive left MCA infarct, involving much of left basal ganglia and cortical gray matter.
P8	75	R	18	123	M	No acute hemorrhage or mass effect identified on early computerized tomography scans. Left basal ganglia calcification present. No other information provided.
P9	72	R	13	23	F	Left MCA territory infarct, including posterior temporal and parietal lobes.
P10	63	R	16	53	M	Left MCA infarct, involving anterior left parietal lobe extending to the level of the sylvan fissure, possible extension into the left temporal lobe.
P11	91	R	18	7	F	Left superior temporal and anterior parietal infarcts with petechial hemorrhage in the left superior temporal a lobe.
<i>Avg (SD)</i>	65.6 (10.8)	12 Right	15.2 (2.0)	67.6 (36.2)	5 M, 6 F	

*Note.* MPO = months postonset of stroke; MCA = middle cerebral artery; R = right; M = male; F = female.

## Appendix B Stimuli

### List A: Blocked condition

blessing  
exchange  
freedom  
merchant  
platform  
announcement  
consequence  
detective  
procession  
sympathy  
civilization  
continuation  
inefficiency  
investigation  
undergraduate

### List B: Random condition

aspect  
creature  
discharge  
garment  
silence  
compliment  
destruction  
employment  
photograph  
preference  
administration  
consideration  
electricity  
individual  
justification

## Appendix C

### Outcome Measures: Definitions and Calculations

Term	Definition
	Two measures:
Error consistency	<ol style="list-style-type: none"> <li>1. Percent consistency of error location</li> <li>2. Percent variability of error type</li> </ol>
1. Consistency of error location	Degree to which sound errors occur in the same target sound three or more times across trials Calculated by dividing the number of sounds consistently in error (within a word type) by the number of total sounds in error
2. Variability of error type	Degree to which sound errors differ from each other within the same location of a word Calculated by dividing the number of errors that differ from each other within the same location of a target word across trials by the total number of errors produced within the same location of the target word across trials
Error rate	Mean number of sound errors per word for each participant
Error type	Number of specified error types (six types) produced by each participant
1. Distortions	An attempt at the target phoneme that does not cross the phoneme boundary, produced with perceptible place, timing, manner, or voice deviation from accurate production
2. Distorted substitutions	A production that not only crosses phoneme boundaries of the target phoneme, but is also distorted
3. Distorted additions	An inserted nontarget phoneme that is distorted
4. Sequential substitutions	Phonemic perseveration, anticipation, and transposition errors
5. Nonsequential substitutions	Phonemic errors other than sequential errors, influenced by factors outside the target word
6. Addition	An inserted nontarget phoneme that is phonetically accurate (not distorted)
7. Omission	A deleted phoneme

## Appendix D

### Error Type Categories and Definitions

Error Type	Definition
Distortion	An attempt at the target phoneme that did not cross the phoneme boundary, produced with perceptible place, timing, manner, or voice deviation from accurate production
Distorted substitution	A production that not only crossed phoneme boundaries of the target phoneme, but was also distorted (van der Merwe, 2009)
Distorted addition	An inserted nontarget phoneme that is distorted
Sequential substitution	Inclusion of phonemic perseveration, anticipation, and transposition errors; these errors are influenced by the context in which the word is produced
Nonsequential substitution	Phonemic errors influenced by factors outside the target word (Mackay & James, 2004)
Addition	An inserted nontarget phoneme that is phonetically accurate
Omission	A deleted phoneme that may result from breakdown at the level of phonological retrieval or motor planning (Buchwald & Miozzo, 2012; Buckingham, 1986; Dell, 1988)

**Appendix E.**

Descriptive Statistics for Percent Consistency of Error Location and Variability of Error type in the Blocked Condition

<b>Descriptive Statistics</b>	<b>Percent Consistency of Error Location</b>		<b>Percent Variability of Error Type</b>	
	Group A	Group P	Group A	Group P
<b>Participants</b>	10	11	10	11
<b>Mean</b>	69.45	71.73	44.31	26.32
<b>Median</b>	69.44	78.11	47.38	21.32
<b>Standard Deviation</b>	11.35	19.59	12.89	14.96
<b>Variance</b>	128.93	383.60	166.08	223.728
<b>Minimum</b>	51.66	29.62	16.67	10.00
<b>Maximum</b>	87.84	100.00	60.56	51.67
<b>Range</b>	36.18	70.38	43.89	41.67
<b>Interquartile Range</b>	15.28	30.53	13.73	26.50
<b>Skewness</b>	-.205	-.964	-1.193	.978
<b>Kurtosis</b>	-.316	.950	1.440	-.537

Descriptive Statistics for Error Rate Group A and Group P

<b>Descriptive Statistics</b>	<b>Error Rate</b>	
	Group A	Group P
<b>Participants</b>	20	22
<b>Mean</b>	2.16	2.48
<b>Median</b>	1.67	1.67
<b>Standard Deviation</b>	1.43	2.29
<b>Variance</b>	2.05	5.22
<b>Minimum</b>	.48	.25
<b>Maximum</b>	5.35	9.40
<b>Range</b>	4.87	9.15
<b>Interquartile Range</b>	1.66	2.81
<b>Skewness</b>	1.085	1.512
<b>Kurtosis</b>	.099	2.629

---

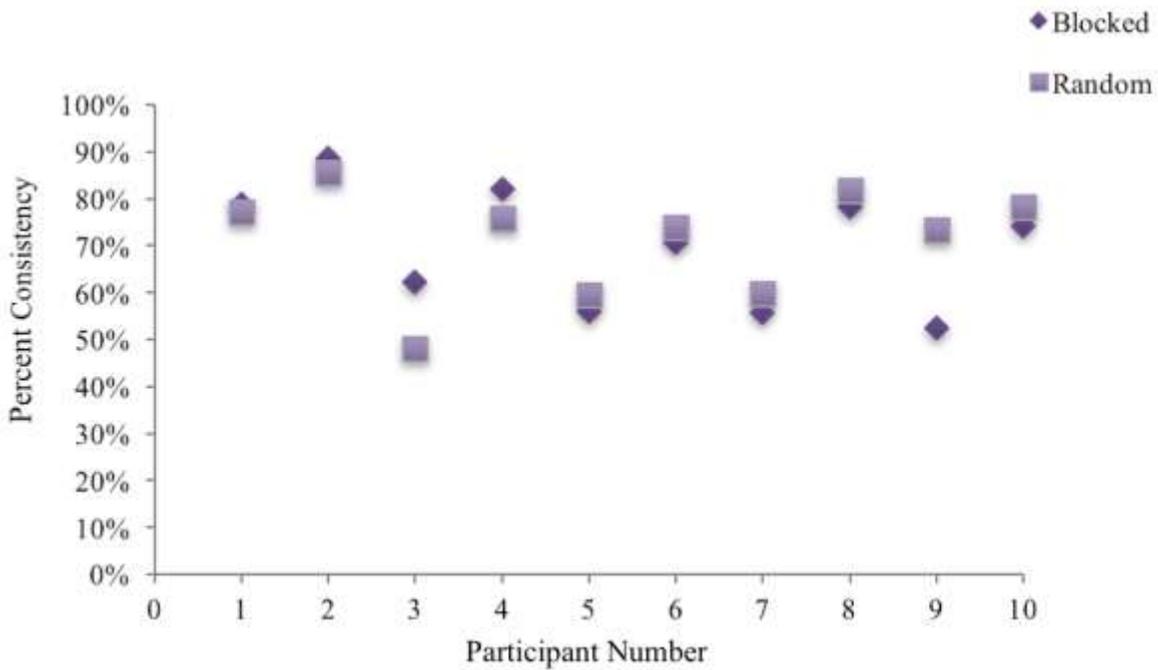
Descriptive Statistics for Severity of Impairment and Error Consistency Across Conditions for Group A and Group P

Descriptive Statistics	Percent Consistency of Error Location					Percent Variability of Error Type				
	Group A			Group P		Group A			Group P	
	Mild	Moderate	Severe	Mild	Moderate	Mild	Moderate	Severe	Mild	Moderate
<b>Participants</b>	6	8	6	12	10	6	8	6	12	10
<b>Mean</b>	72.98	69.21	64.45	72.20	65.25	35.73	41.05	47.02	24.58	43.18
<b>Median</b>	76.18	70.68	65.55	73.39	69.68	33.78	41.43	48.36	20.66	51.57
<b>Standard Deviation</b>	15.59	11.07	4.51	18.12	18.46	22.09	6.67	4.34	17.48	21.62
<b>Variance</b>	243.18	122.57	20.33	328.22	340.86	488.07	44.48	18.80	305.527	467.34
<b>Minimum</b>	53.00	51.66	55.74	27.50	29.62	13.33	29.01	38.62	.00	15.17
<b>Maximum</b>	95.00	87.84	68.58	100.00	89.57	60.56	48.94	50.35	65.42	77.84
<b>Range</b>	42.00	36.18	12.84	72.50	59.95	47.23	19.93	11.73	65.42	62.67
<b>Interquartile Range</b>	28.10	14.67	4.84	19.71	27.28	41.49	10.64	5.50	24.11	36.65
<b>Skewness</b>	-.068	-.077	-1.869	-1.205	-.620	.075	-.651	-1.939	1.125	-.059
<b>Kurtosis</b>	-.771	.585	4.401	2.875	-.177	-3.036	-.042	4.021	1.660	-1.338

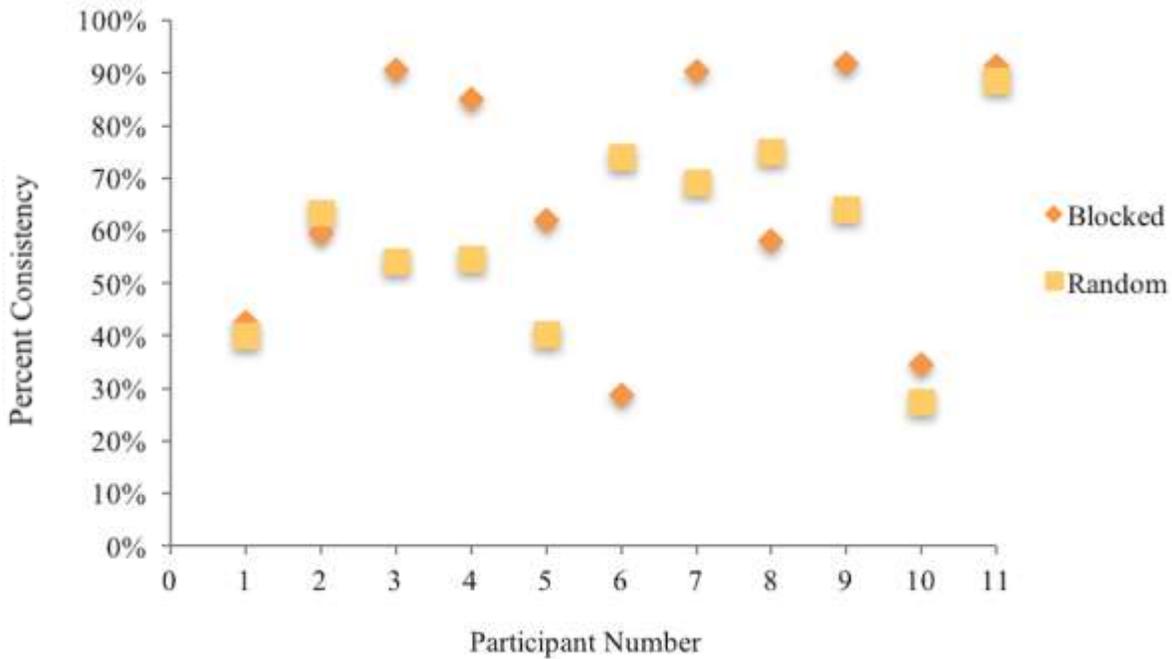
Descriptive Statistics for Group Differences on Error Type in the Blocked Condition

Descriptive Statistics	<u>Distortions</u>		<u>Distorted Substitutions</u>		<u>Distorted Additions</u>		<u>Substitutions</u>		<u>Additions</u>		<u>Omissions</u>	
	Group A	Group P	Group A	Group P	Group A	Group P	Group A	Group P	Group A	Group P	Group A	Group P
<u>Participant</u>	10	11	10	11	10	11	10	11	10	11	10	11
<u>Mean</u>	76.30	33.27	29.40	9.55	29.10	7.45	81.40	80.55	46.80	47.18	38.90	61.55
<u>Median</u>	70.00	35.00	20.50	10.00	25.00	6.00	58.00	55.00	31.00	26.00	19.00	50
<u>Standard Deviation</u>	29.22	12.09	19.96	6.09	17.48	5.65	68.67	70.83	41.81	51.08	45.53	51.09
<u>Variance</u>	853.78	146.21	398.49	37.07	305.43	31.87	4715.60	5017.07	1747.96	2609.96	2072.77	2609.67
<u>Minimum</u>	39	13	11	1	8	1	17	14	5	0	3	0
<u>Maximum</u>	120	54	65	23	55	18	231	222	143	139	139	135
<u>Range</u>	81	41	54	22	47	17	214	208	138	139	136	135
<u>Interquartile Range</u>	60	14	30	8	34	8	95	113	32	93	67	110
<u>Skewness</u>	.338	-.316	1.164	.772	.491	.661	1.406	1.061	1.746	-1.143	1.498	.305
<u>Kurtosis</u>	-1.435	.179	-.013	1.367	-1.118	-.744	1.343	.065	2.589	-.384	1.412	-1.653

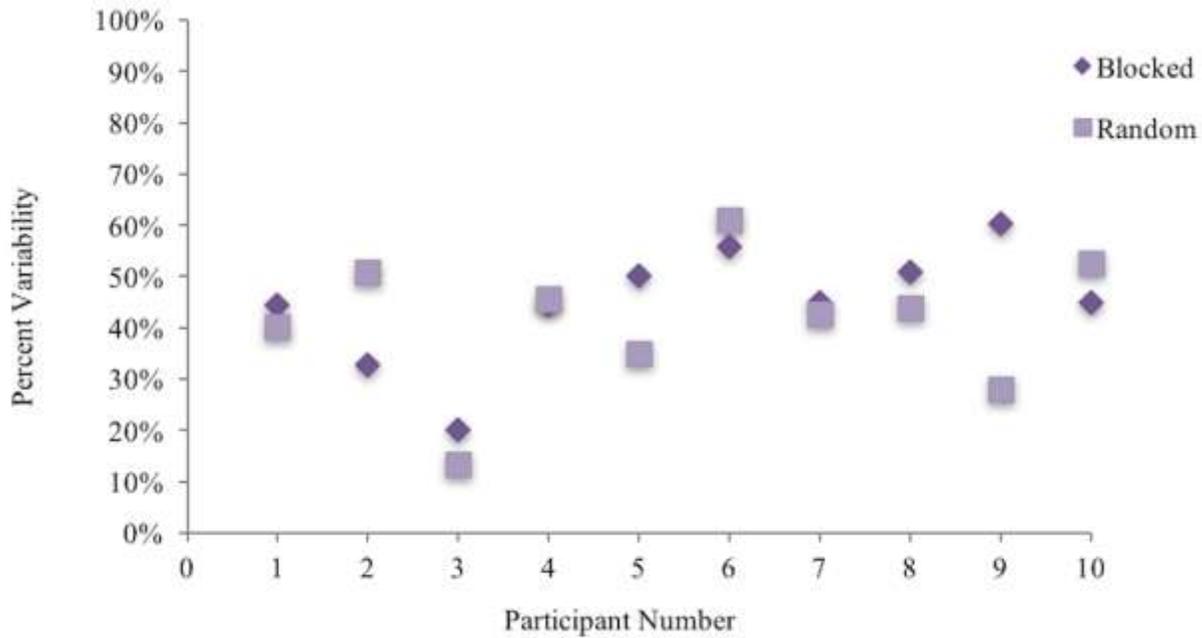
## Appendix F



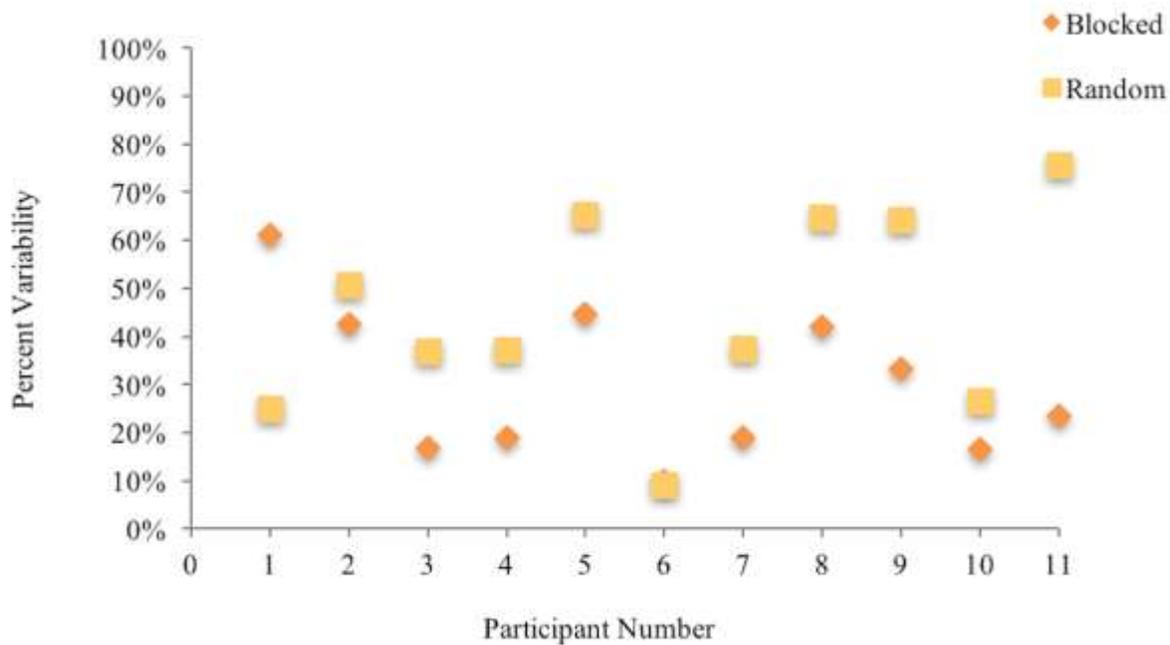
Group A: Individual Performance on Consistency of Error Location in Blocked and Random Presentation Conditions



Group P: Individual Performance on Consistency of Error Location in Blocked and Random Presentation Condition



Group A: Individual Performance on Variability of Error Type in Blocked and Random Presentation Conditions



Group P: Individual Performance on Variability of Error Type in Blocked and Random Presentation Conditions