Research Article

Identification and Remediation of Phonological and Motor Errors in Acquired Sound Production Impairment

Adam Buchwald,^a Bernadine Gagnon,^b and Michele Miozzo^{c,d}

Purpose: This study aimed to test whether an approach to distinguishing errors arising in phonological processing from those arising in motor planning also predicts the extent to which repetition-based training can lead to improved production of difficult sound sequences.

Method: Four individuals with acquired speech production impairment who produced consonant cluster errors involving deletion were examined using a repetition task. We compared the acoustic details of productions with deletion errors in target consonant clusters to singleton consonants. Changes in accuracy over the course of the study were also compared.

Results: Two individuals produced deletion errors consistent with a phonological locus of the errors, and 2 individuals produced errors consistent with a motoric locus of the errors. The 2 individuals who made phonologically driven errors showed no change in performance on a repetition training task, whereas the 2 individuals with motoric errors improved in their production of both trained and untrained items.

Conclusions: The results extend previous findings about a metric for identifying the source of sound production errors in individuals with both apraxia of speech and aphasia. In particular, this work may provide a tool for identifying predominant error types in individuals with complex deficits.

This special issue contains selected papers from the March 2016 Conference on Motor Speech held in Newport Beach, CA.

Understanding the distinction between phonological processing and motor planning has been a persistent issue in both the clinical and neurocognitive literatures on speech and language production. The issue has persisted both because of its importance and its difficulty. From a clinical perspective, this issue has primarily been addressed in the literature on the differential diagnosis of apraxia of speech (AOS) and phonemic paraphasia (McNeil, Robin, & Schmidt, 2009; Wambaugh, Duffy, McNeil, Robin, & Rogers, 2006a) and in studies on the treatment of these deficits (Ballard et al., 2015; Wambaugh, Duffy, McNeil, Robin, & Rogers, 2006b). From a neurocognitive perspective, the focus has been on identifying the mechanisms involved in phonological processing and motor planning (Buchwald &

Miozzo, 2011, 2012; Galluzzi, Bureca, Guariglia, & Romani, 2015; Laganaro, 2012; Levelt, Roelofs, & Meyer, 1999; Ziegler & Aichert, 2015; Ziegler, Staiger, & Aichert, 2010).

One key difficulty in addressing this issue is the overlapping nature of the phonological and motor systems and their related impairments. The presence of “pure” AOS in the absence of other impairments is described as rare (McNeil et al., 2009). Therefore, a number of investigations of speakers with AOS with concomitant aphasic impairments attempt to isolate the motor control system by using tasks that likely tap directly into this system (Haley, Jacks, & Cunningham, 2013; Maas, Gutiérrez, & Ballard, 2014; Maas, Mailend, & Guenther, 2015). A related issue is that the types of sound structure sequences that are difficult for the motor system to plan and enact (e.g., consonant clusters) are precisely those that are more complex phonologically as well. It has been argued that the phonological complexity leads them to be acquired later, avoided cross-linguistically, and sensitive to acquired deficits (Jakobson, 1941/1968).

Our approach to this issue has been built on the foundational work of distinguishing motor planning impairment in AOS from phonological impairment in aphasia. However, because of the high rate of comorbidity of these impairments, we have concentrated on the errors observed in both of these deficits. We consider these errors within the framework of neurocognitive accounts of speech and language

^aNew York University

^bTeachers College, New York

^cThe New School for Social Science Research, New York

^dJohns Hopkins University, Baltimore, MD

Correspondence to Adam Buchwald: buchwald@nyu.edu

Editor: Yana Yunusova

Associate Editor: Kristie Spencer

Received June 15, 2016

Revision received October 25, 2016

Accepted November 24, 2016

https://doi.org/10.1044/2017_JSLHR-S-16-0240

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

production, and we have examined their acoustic and articulatory properties to identify the level of processing where an error arises (insertion errors: Buchwald et al., 2007; deletion errors: Buchwald & Miozzo, 2011, 2012). Although identifying the level in which errors arise is a notable challenge in individuals who present with both aphasia and AOS, it is also especially important because this represents the majority of individuals with AOS. At this point, we believe that our approach allows us to—at minimum—determine the typical error types produced by an individual.

Our approach relies on the widely shared hypothesis that there are interacting but distinct levels of processing involved in spoken word production (e.g., Dell, Schwartz, Martin, Saffran, & Gagnon, 1997; Rapp & Goldrick, 2000), including levels involved in phonological processing and motor planning. Furthermore, errors can arise within each of these systems. At the phonological processing level, speakers retrieve and encode phonological representations from long-term memory. Although there are differing views about the content of these representations (e.g., abstract phonemes in Dell, 1986, 1988 et seq. vs. articulatory gestures in Browman & Goldstein, 1986, 1990 et seq.), there is a shared idea that speakers store context-independent information that is then mapped—based on language-specific phonology—onto a detailed phonological representation whose content is also described in various ways in the literature (e.g., allophones vs. gestural score). This more detailed phonological representation is mapped onto a series of motor plans that are executed by the speech musculature.

There is also debate about the degree of interaction among these levels. Whereas some foundational views posited a modular system in which one level becomes active only after the previous one has been completed (Levelt et al., 1999), there is ample evidence that the interaction between levels takes the form of simultaneous activation in spoken production (e.g., Dell et al., 1997; Rapp & Goldrick, 2000; Laganaro, 2012; Morsella & Miozzo, 2002), sign language production (Navarrete, Peressotti, Lerose, & Miozzo, 2017), and written production (Buchwald & Falconer, 2014; Falconer & Buchwald, 2013; McCloskey, Macaruso, & Rapp, 2006). However, although there is spreading activation and thus interaction in these frameworks, they retain the property of being “staged” where a final computation at one level does not conclude until the previous level has concluded (Dell et al., 1997; Nozari, Dell, & Schwartz, 2011; Nozari, Kittredge, Dell, & Schwartz, 2010). Thus, although the presence of simultaneous activation in these systems is different from a traditional view of cognitive modules (Fodor, 1983) that are informationally encapsulated and automatic, these accounts still posit levels of neurocognitive processing that compute different aspects of the speech representation. We return to a discussion of interactivity in the speech and language production system in the General Discussion.

Acquired Sound Production Impairment

A clear distinction has been proposed in the literature on acquired sound production impairment between AOS and aphasic deficits that affect phonological processing.

Despite early descriptions of AOS as having a phonological component (Darley, Aronson, & Brown, 1975; Wertz, Lapointe, & Rosenbek, 1984), it is now typically held that AOS is, by definition, a motor speech impairment and is thus qualitatively different from impairments affecting phonological processing (McNeil, Pratt, & Fossett, 2004; van der Merwe, 2009; Wambaugh et al., 2006a). Phonological impairments have often been observed in individuals with conduction aphasia and have been described as aphasia with phonemic paraphasia in the literature on differential diagnosis (e.g., McNeil et al., 2009). Consistent with the neurocognitive approach adopted here, we will use the term *phonological aphasia* (PA) in the remainder of this article to refer to acquired phonological impairment, encompassing the broad group of aphasia subtypes that can affect phonological processing.

Evidence in support of the distinctive basis of AOS and PA emerged from case reports of AOS without concomitant aphasia (McNeil, Odell, Miller, & Hunter, 1995; Odell, McNeil, Rosenbek, & Hunter, 1990, 1991). These reports were instrumental in defining the differential diagnosis of PA and AOS, identifying slow, dysprosodic speech with sound distortions and perceived sound substitutions as hallmarks of AOS (McNeil et al., 2009; Wambaugh et al., 2006a). The dysprosodic component is marked by syllable segregation and a lack of appropriate stress marking often described as equal stress across adjacent syllables. An additional marker—error consistency—had originally been identified as a core feature of AOS (see review in Wambaugh et al., 2006a), but recent evidence has called this feature into question (Haley et al., 2013; Staiger, Finger-Berg, Aichert, & Ziegler, 2012). The specific deficits observed in AOS are generally attributed to impairment to the motor planning system, particularly in the spatial and temporal planning/programming of articulatory movements. For both PA and AOS, it is rare that these deficits occur in isolation in the absence of other impairment (Duffy, 2005).

As highlighted in two systematic reviews of studies on AOS treatment that together cover all articles published from 1951 to 2012 (Ballard et al., 2015; Wambaugh et al., 2006b), both articulatory–kinematic and rate/rhythm approaches to AOS treatment can be effective in rehabilitation, with the preponderance of available evidence supporting articulatory–kinematic approaches to facilitate speech–sound production. In addition, there is evidence that these approaches are effective when practice and feedback are structured in accordance with general principles of motor learning (Bislick, Weir, Spencer, Kendall, & Yorkston, 2012; Maas et al., 2008; Wambaugh, Mauszycki, & Ballard, 2013; Wambaugh, Nessler, Wright, & Mauszycki, 2014; Wambaugh, Nessler, Wright, Mauszycki, & DeLong, 2016). Our approach discussed below is consistent with these results, as we find a decrease in motor planning errors following repetition-based practice structured according to principles of motor learning.

Neurocognitive Approaches to Sound Production

Neurocognitive accounts have been built on an information processing view of speech production that assumes

a series of processes that are ordered (e.g., Levelt et al., 1999; but see Miozzo, Pulvermüller, & Hauk, 2015) and interact with one another via spreading activation (e.g., Dell et al., 1997; Morsella & Miozzo, 2002; Rapp & Goldrick, 2000). The presence of interaction has been argued to account for a variety of findings, including effects of the structure of the lexicon on the details of speech production as well as the presence of traces of target utterances in speech errors (Kurowski & Blumstein, 2016). However, the presence of interaction does not at all imply that there are not distinct processes involved in production (see Rapp & Goldrick, 2000, for a discussion of limited interaction), and our present approach capitalizes on a classic distinction within this literature. These accounts provide a framework for identifying markers of speech-production deficits that would reveal what mechanisms are specifically impaired. Ziegler and colleagues (e.g., Ziegler et al., 2010) focused on syllable frequency, a variable reflecting the retrieval from a mental syllabary of a motor plan corresponding to a syllable that is ultimately produced. Consistent with the production architecture outlined by Levelt et al. (1999), access to the mental syllabary follows syllabification that takes place during phonological processing; in addition, it is not sensitive to syllable frequency. Evidence that individuals with AOS are more accurate in producing high- versus low-frequency syllables (German: Aichert & Ziegler, 2004; Staiger & Ziegler, 2008; French: Laganaro, 2005, 2008) suggests that AOS affects mechanisms supporting motor planning rather than phonological processing.

In the present study, we followed traditional neurocognitive accounts (Dell, 1986; Dell et al., 1997; Goldrick & Rapp, 2007; Levelt et al., 1999; Rapp & Goldrick, 2000; Stemmer, 1985) in assuming that speech production minimally involves (a) accessing and encoding a word's phonological representation from long-term memory, (b) mapping from this representation to a context-specific representation that encodes position-specific sound properties, and (c) generating a motor plan on the basis of this context-specific representation. Although there has been rich debate about the precise nature of these representations and processes (Aichert & Ziegler, 2004; Browman & Goldstein, 1990; Chomsky & Halle, 1968; Goldrick & Rapp, 2007; Guenther & Perkell, 2004; Prince & Smolensky, 1993/2004; Romani, Galluzzi, Bureca, & Olson, 2011; van der Merwe, 2009; Varley & Whiteside, 2001), the present research relies on one issue with relatively widespread agreement: the existence of a phonological level of processing that occurs before context-specific representations are generated and fully specified, and a motor planning level of processing that occurs after context-specific representations are generated (Dell et al., 1997; Goldrick & Rapp, 2007; Goldstein, Byrd, & Saltzman, 2006; Levelt et al., 1999).

In our previous investigations (Buchwald & Miozzo, 2011, 2012; Miozzo & Buchwald, 2013), we presented data from two individuals who presented with similar clinical profiles that are characteristic of co-occurring aphasia and AOS, and who made similar sound production errors in speech production (deletion of /s/ in /s/-initial consonant

clusters). However, a clear difference was observed in their errors involving /s/-deletion (e.g., /spik/ → [pik]; Buchwald & Miozzo, 2011). One individual tended to produce the resulting stop (e.g., the /p/ in [pik]) with aspiration, as speakers would typically produce a word-initial voiceless singleton (and similarly to how that speaker produced the /p/). This is consistent with a deletion occurring at a phonological level before the context-specific representations are computed. We refer to deletion errors in which the resulting sound is produced with singleton timing as *phonological errors* throughout the remainder of this article. In contrast, the other individual tended to produce the resulting stop without aspiration, as speakers typically produced the stop in a consonant cluster (which is acoustically similar to a voiced singleton stop in American English). This is consistent with an error that occurs during motor planning, after the presence of an onset consonant cluster has been specified. We refer to errors in which the resulting sound is produced with cluster timing as *motor errors* throughout the remainder of this article. In a follow-up article (Buchwald & Miozzo, 2012), we also observed that the pattern of producing cluster timing versus singleton timing was maintained for /s/-deletion in other /s/-initial clusters (e.g., s-nasal onsets; /snæk/ → [næk]). In this case, the individual whose errors we identified as phonological produced the resulting nasal with timing of a singleton (longer than the nasal in an s-nasal onset), whereas the individual whose errors we identified as motoric produced a nasal with the timing of a nasal in an s-nasal onset cluster, which is shorter than the singleton nasal.

In addition to the distinction in the quality of the errors, we found clear differences between the two individuals we tested with respect to factors that led to fewer errors. In particular, in a notable post hoc finding, we observed that the individual whose errors arose at a motoric level showed a marked decrease in error rates over the course of 6 months of testing. In contrast, the individual who made primarily phonological errors did not show a similar improvement. Each of these individuals was administered repetition tasks weekly, which included the clusters described above. The repetition sessions were distributed over several months, resulting in a large amount of practice, and words containing consonant clusters were presented at a random schedule with variable targets. The distribution, amount of practice, schedule, and targets reflect four parameters of the structure of practice sessions that are consistent with the principles of motor learning (Maas et al., 2008). It is also worth noting that the speakers did not receive feedback, so these parameters focus on the structure of practice. Thus, in considering the results from Buchwald and Miozzo (2012), we hypothesized that the benefit for the individual with motoric errors arose because the repetition task functioned as a speech motor learning intervention. However, the limited number of participants, the limited number of clusters that were investigated, and the post hoc and uncontrolled nature of the finding of improvement necessitated a follow-up study in which we addressed these limitations. The present study reflects an important step in that endeavor.

Overview of the Present Study

The purpose of the present study was to examine the relationship between error types and improvement in cluster production following repetition-based speech motor learning practice sessions in four individuals with a diagnosis of aphasia and AOS. Generalization of trained clusters from trained items to untrained items was also examined. The following research questions were addressed:

1. Can we use deletion errors in clusters to detect differences in error types in individuals with co-occurring aphasia and AOS?
2. Do individuals with predominantly motor errors improve in their cluster production following extensive exposure to words containing those clusters? Do individuals with predominantly phonological errors improve?
3. Do improvements in cluster production generalize from words used in training to untrained words that also contain those clusters?

The hypotheses for this study were based on our previous work (Buchwald & Miozzo, 2011, 2012) as well as on theoretical support from motor learning theory (Bislick et al., 2012; Maas et al., 2008; Schmidt & Lee, 2005). We hypothesized that the participants whose errors were identified as arising during motoric processing would show an improvement in cluster accuracy between the pretraining and posttraining sessions as a result of training sessions that promote speech motor learning in these participants. We also predicted that the learning would occur at the level of clusters rather than lexical items, so the benefit should extend to untrained lexical items that begin with the same consonant clusters. In contrast, we predicted that the participants whose errors were identified as arising during phonological processing would not show a change in accuracy as a result of training sessions that promote speech motor learning because this approach would not address the underlying deficit.

Method

Participants

The study included four participants (P1, P2, P3, and P4) with suspected AOS and aphasia. The participants were all presented and passed a hearing screening at 500, 1000, 2000, and 4000 Hz at 40 dB in at least one ear. Following other researchers (e.g., Maas et al., 2015), we used 40 dB to minimize the possibility that the older participants would be excluded. All testing was conducted in a sound-treated (but not soundproof) room. All procedures were approved by the institutional review board (IRB) at the place of testing and informed consent. Procedures for P1 and P3 were approved by the IRB at the Teacher's College, where they provided informed consent; testing for P2 and P4 was approved by the IRB at New York University, where these individuals provided informed consent.

The diagnosis of AOS (or suspected AOS) primarily followed the guidelines in the research literature (Wambaugh et al., 2006a). The one exception is that we did not consider error consistency as a criterion following the debate described in the Introduction. Two certified speech-language pathologists with experience diagnosing motor speech disorders independently rated each participant on a 3-point scale (following Maas et al., 2015): 1 = *no AOS*; 2 = *possible AOS*; 3 = *AOS*. These ratings were based on a variety of speech samples, including those collected as part of standardized tests (e.g., the Western Aphasia Battery–Revised [WAB-R]; Kertesz, 2006), as well as descriptions of other pictures and a spontaneous speech sample. The diagnosis of AOS focused on slowed speech with segmental and inter-segmental prolongations, dysprosody, and distortions and distorted substitutions. The lack of dysprosody would have automatically led to a rating of 1. In addition to displaying these primary characteristics, all participants exhibited some nondiscriminative behaviors that were not used to diagnose AOS: articulatory groping, difficulty with initiation, more errors on longer words, and occasional islands of fluent speech. In addition to the clinician ratings, the first author independently rated each participant on the basis of his experience as an AOS researcher.

Of the four participants, the three ratings were in agreement on P2 (all rated 3) and P4 (all rated 2). All remaining participants were rated as either 2 or 3 by each independent rater. Table 1 presents demographic data about each individual as well as their performance on naming, articulation, motor speech, and auditory discrimination tasks when available. Aphasia type and severity were assessed using the WAB-R for P2, P3, and P4 and are reported in Table 1. The whole WAB-R was unfortunately not available for P1, and we were unable to compute an aphasia quotient (AQ) score for that participant. In addition to Table 1, we provide a brief overview of each participant here and include information about the language abilities of P1 from other tests. All participants were native speakers of American English.

P1 is a right-handed man. A neurologist's report indicated that he suffered a single left middle cerebral artery (MCA) cerebrovascular accident (CVA), leaving him with little to no hemiplegia or weakness. We did not obtain an AQ for P1. His speech was rated as 4 according to the WAB fluency scoring system in describing the Cookie Theft picture. He was administered two subtests of the Boston Diagnostic Aphasia Examination–Third Edition (Goodglass, Kaplan, & Barresi, 2001) to evaluate auditory comprehension. On the commands subtest, he performed well on one- and two-step commands but showed difficulty with multistep commands. On the Complex Ideational Material test, he showed difficulty with syntactically complex yes/no questions (2/4). His repetition was impaired, particularly with longer phrases. His picture naming and word finding also revealed impairment, scoring 35/60 on the Boston Naming Test (Kaplan, Goodglass, & Weintraub, 1983) pictures and naming five animals within 1 min on the category naming task.

Table 1. Demographic information and results from standardized tests.

Parameter	P1	P2	P3	P4
Age (years)	50	60	44	66
Sex	M	M	F	M
Education (years)	14	21	18	20
Time postonset (months)	84	56	60	30
Handedness	Right	Right	Right	Right
Etiology	LH CVA	LH CVA	LH CVA	LH CVA
Naming				
BNT	35/60	1/15 (short)	15/28 (subset)	
PNT				128/175
AOS rating	2.7	3.0	2.7	2.0
Auditory discrimination (PALPA 2)	69/72	72/72	69/72	67/72
Dysarthria	Mild	None	None	None
WAB-R AQ score		44.1	84.2	83.8
Category	(Broca's)	Broca's	Anomic	Conduction

Note. LH CVA = left-hemisphere cerebrovascular accident. The naming results come from the Boston Naming Test (BNT) and the Philadelphia Naming Test (PNT; Roach, Schwartz, Martin, Grewal, & Brecher, 1996). The apraxia of speech (AOS) rating was the mean rating across three independent raters. Auditory discrimination scores come from the minimal pair discrimination test of the Psycholinguistic Assessments of Language Processing in Aphasia (PALPA 2). Dysarthria was evaluated on the basis of an oral motor examination (Duffy, 2005), and the aphasia type and aphasia quotient (AQ) score came from the Western Aphasia Battery–Revised (WAB-R). For P1, the aphasia type is based on his performance on tests similar to those in the WAB-R, which are discussed in the text.

Overall, his profile is consistent with a diagnosis of Broca's aphasia.

P2 is a right-handed man. A neurologist's report indicated that he suffered a single left MCA infarct, leaving him with right-sided hemiplegia. Brain scans were not available. P2's AQ was 44.1 and consistent with Broca's aphasia (nonfluent output, receptive language largely intact, impaired repetition and naming).

P3 is a right-handed woman. A neurologist's report indicated that she suffered a single left MCA infarct, with the lesion primarily seen in frontoparietal regions, leaving her with mild right-sided hemiplegia. P3's AQ was 84.2, and her performance was consistent with a diagnosis of anomia (fluent, intact comprehension, good repetition, impaired naming and word finding).

P4 is a right-handed man. A neurologist's report indicated a left MCA infarct affecting the posterior portion of the frontal lobe and the parietal lobe on the left side, leaving him with mild right-sided weakness. A secondary infarct was present in the right frontal lobe; it was not known when the secondary infarct occurred. P4's AQ was 83.8 and consistent with Broca's aphasia (fluent output, receptive language largely intact, impaired repetition, and mild naming and word finding impairment).

Auditory Perception

In addition to the pure tone hearing screening described above, we extensively probed participants' auditory comprehension using minimal pair discrimination tasks from the Psycholinguistic Assessments of Language Processing in Aphasia (PALPA; Kay, Lesser, & Coltheart, 1992) involving words (PALPA 2). This was done in order to verify that errors in repetition came from an articulatory

deficit rather than an auditory word recognition deficit. Word pairs were spoken one at the time by the experimenter, and participants made a same/different decision task. Same and different pairs occurred with identical probabilities (.5 each). Participants performed within controls' range in a standardized task from.

Procedure and Data Analysis

Repetition Training Task

Each of the participants performed the same training, which consisted of eight weekly sessions of a repetition task with 131 or 188 words (see below). Each session lasted between 30 and 50 min. In the repetition task, the experimenter said the word once, and the participant was asked to repeat it. The first full response was scored. No feedback was provided to the participant. The next word was administered within 5 s of the end of the response. The repetition task consisted of monosyllabic words containing consonant clusters, with a total of 12 clusters being tested (stop-liquid: /br/, /pl/, /gl/, /kr/; /s/-stop: /spl/, /st/, /sk/; s-nasal: /sm/, /sn/; fricative-liquid: /fl/, /fr/, /sl/). These clusters were selected because deletion errors would be conducive to the acoustic analyses performed in this article, as described below. Word lists were randomized within and across participants during each training session. In addition to the repetition training, each participant was administered a larger list of words as a pretest and a posttest. These larger lists were administered in the week(s) immediately before the first session and immediately after the last session. This set included trained words, untrained words matched for onset cluster and frequency, and additional filler items that were not part of the analysis. There were a total of 262 words with these critical consonant

clusters. For P1 and P3, 131 of the items were trained, and 131 were untrained. Owing to experimental error, for P2 and P4, 188 items were used in the weekly training, with 74 remaining untrained. P1 and P3 performed the pre- and posttests once each; P2 and P4 performed them twice each. We will return to this issue in the Limitations and Future Directions section of the General Discussion.

In addition, for participants P1, P2, and P3, we were able to bring them back for additional testing within a month of the posttest during which we administered an additional repetition task. This task contained singleton words as controls for the deletion errors (details below) as part of a larger list of monosyllabic words (as described above). This task was performed by the same experimenter and in the same way as the other testing. P4 was unable to return.

Acoustic Measures of Deletion Errors

Across the pretests, training sessions, and posttest, each participant produced 3,008 stimulus words. To determine whether the deletion errors (e.g., *smear* → *smear*) reflected the articulatory timing of singletons (as predicted for phonological processing errors) or clusters (as predicted for motor planning errors), we identified each deletion error in consonant clusters in both onset and coda clusters within this set of stimuli. Deletion errors were identified by transcription with simultaneous examination of the acoustic record. An item was identified as containing a deletion only when there was no trace of the sound in the spectrogram or waveform, and the sound was perceived to be deleted as well. For these items, if the original online transcription did not reflect a deletion, a third transcriber resolved the discrepancy. Thus, 100% of the items identified as deletions were agreed upon by multiple independent transcribers. All onset errors containing deletion of the initial consonant with a nasal or liquid as the second consonant (e.g., *small*, *plane*) were used in the analysis. These items were selected for analysis because the remaining sound has an intrinsic duration, and the acoustic duration of these sounds has been reported to be shorter when in clusters compared to singletons (O'Shaughnessy, 1974). For coda clusters, the clearest shortening in clusters relative to singletons reported by O'Shaughnessy (1974) arose in the nasal portion of homorganic nasal-stop codas (e.g., /mp/, /nt/). Therefore, our analyses also included tokens with deletion of stop consonants in these clusters. This allowed us to retain a consistent relationship in all the errors we analyzed such that the length of a segment in the cluster was expected to be shorter than that same segment as a singleton.

One benefit of this approach is that each stimulus word was attempted at least four times (for items in the pre- and posttest only) and as many as 12 times (for trained items). Thus, we were consistently able to compare the duration of segments in tokens containing deletions to the duration of the same segment in accurately produced tokens (e.g., compare the /m/ in *smear* → [mir] with the /m/ in accurately produced *smear* → [smir]). We selected the

immediately subsequent accurate production (i.e., from the next week that the participant produced the word correctly) to compare with the deletion. For errors that arise in phonological processing, we expect the segment in the token with the deletion error to be systematically longer than in the accurately produced token. Contrariwise, we expect these two durations to be equivalent for errors that arise after the context-dependent plans have been specified. For these tokens especially, we elicited matched singletons (e.g., *mere*) because we expect the singleton to be longer than the deletion for errors arising in motor processing. This was done after the main data collection process in order to target the individual control words required for this analysis. We intended to perform this analysis for all individuals; however, P4 was unable to participate following the posttest, so we collected singletons only for P1, P2, and P3. We will address this limitation in the General Discussion.

The measurements were made by phonetically trained research assistants using information from the waveform and the spectrogram of Praat (Boersma & Weenink, 2015). In addition, 20% of the productions for each participant (including errors and control tokens) were remeasured by the first author. The measures were within 5 ms of one another for all but five of the measures (5/81; 6.2%) spread across the four participants. We also note that by focusing on deletion errors, our acoustic analyses were performed on a relatively small subset of the errors, and we address this issue in the General Discussion as well.

Training Outcome Measures

Our primary training outcome measure was the change in cluster production accuracy for the 12 clusters listed above, reported separately for trained and untrained items. To determine cluster accuracy, each token from each participant was evaluated online in broad transcription and was later transcribed again on the basis of the acoustic recording (by a trained listener who was blind to the original transcription). All cases in which a difference between these scores occurred were resolved by a third listener who listened independently. Thus, the cluster accuracy for all items was agreed upon by multiple transcribers. For this analysis, clusters were identified as correct or incorrect, and all tokens were included regardless of error type. The change in accuracy was the simple subtraction of posttraining accuracy minus pretraining accuracy.

Results

The performance of each of the four individuals was analyzed separately to determine how their error productions compared with their accurate productions and whether they benefited from the repetition training. We present the results of the acoustic analyses first to establish the error types and then present the results of changes that emerged owing to training. The errors examined in the acoustic analyses were produced in each of the study sections (pretest, training task, or posttest).

Acoustic Analysis of Error Types

P1

P1 made 498 errors across testing sessions. Of these, there were 33 deletion errors that fit the criteria described above, each of which occurred in coda clusters involving the deletion of final stops from homorganic nasal-stop clusters (e.g., *scamp* → [skæm]). (None of the onset errors made by P1 were deletion errors.) For each error, we were able to obtain a token with the cluster accurately produced (e.g., *scamp* → [skæmp]) and a correctly produced singleton coda (e.g., *scam* → [kæm]). Because of the difficulty of identifying the offset of a vowel prior to the nasal coda consonant, we measured the overall duration of the vowel+nasal sequence from each token (e.g., the [æm] in each word). These stimuli were analyzed using a one-way analysis of variance (ANOVA) comparing the duration of the vowel+nasal consonant sequence when produced in a cluster, a singleton, and a deletion error. The ANOVA indicated a significant difference among groups, $F(2, 96) = 11.37, p < .001$. Follow-up t tests indicated that the duration of the sequence in tokens with a singleton coda ($M = 485$ ms, $SD = 146$ ms) was significantly longer than the duration in tokens with deletion ($M = 340$ ms, $SD = 146$ ms), $t(32) = 6.09, p < .001$, or the tokens with accurately produced clusters ($M = 333$ ms, $SD = 145$ ms), $t(32) = 6.47, p < .001$, with no difference between cluster and deletion tokens, $t(32) = 0.79, p = .44$. These patterns are consistent with errors arising after the generation of context-dependent articulatory timing, matching the timing associated with a cluster (see Figure 1a).

P2

P2 made 1,464 errors throughout the testing sessions. Of these errors, there were 21 deletion errors in onset clusters with either /t/ or /l/ as C2 (e.g., *grid* → [rid]). For each error, we were able to obtain a token with the cluster accurately produced (e.g., *grid* → [grid]) and a correctly produced singleton onset (e.g., *rid* → [rid]). We measured the overall duration of the liquid+vowel from each of these words. These measures were analyzed using an ANOVA comparing the duration of the liquid+vowel sequence when produced in a cluster, a singleton, and a deletion error. The ANOVA indicated a significant difference among groups, $F(2, 60) = 3.77, p < .03$. Follow-up t tests indicated that the duration of the sequence in tokens with a singleton onset ($M = 355$ ms, $SD = 82$ ms) was significantly longer than the duration in tokens with deletion ($M = 303$ ms, $SD = 54$ ms), $t(20) = 5.75, p < .001$, or accurately produced clusters ($M = 313$ ms, $SD = 90$ ms), $t(20) = 2.91, p < .01$, with no difference between the correct cluster and deletion tokens. These patterns are consistent with errors that arise after the generation of context-dependent articulatory timing (see Figure 1b).

P3

P3 made 446 errors across all testing sessions. Of these, there were 52 deletion errors in onset clusters with either nasal or liquid C2 (e.g., *gland* → [lænd]). For each

error, we were able to obtain a token with the cluster accurately produced (e.g., *gland* → [glænd]) and a correctly produced singleton onset (e.g., *land* → [lænd]). We measured the overall duration of the nasal or liquid from each of these words. The duration of the consonant differed when produced in a cluster, a singleton, and a deletion error, $F(2, 153) = 77.39, p < .001$. Follow-up t tests indicated that the duration of the consonant in tokens with a cluster onset ($M = 75$ ms, $SD = 14$ ms) was significantly shorter than the duration in tokens with deletion ($M = 106$ ms, $SD = 18$ ms), $t(32) = 10.01, p < .001$, or the tokens with accurately produced singletons ($M = 109$ ms, $SD = 14$ ms), $t(51) = 14.21, p < .001$, with no difference between the singleton and the deletion tokens. These patterns are consistent with errors that arise prior to the generation of context-dependent articulatory timing, thus matching the timing associated with a singleton onset (see Figure 1c).

P4

P4 made 302 errors over the course of testing. Of these, there were 35 deletion errors in onset and coda clusters. Twenty of these errors involved /s/-initial clusters in which C2 was deleted, leaving the /s/ in the onset (e.g., *sleep* → [sip]). Previous research has not shown a clear and systematic difference in the acoustic duration of /s/ as a singleton onset versus as the initial consonant of a cluster. Thus, these errors were excluded from our acoustic analysis, leaving 15 clusters with deletion. P4's deletion error contexts were less systematic than those of the other participants and included onset clusters with C1 deletion (e.g., *broke* → [rok]), coda clusters with the second consonant deleted (e.g., *brand* → [bræn]), and s-stop clusters with the /s/ deleted (e.g., *steam* → [tim]) for which we can measure the voice onset time (VOT). For these 15 words with deletion, we identified the segment in a correctly produced cluster for comparison (e.g., *d* in *brand* → [brænd]). In the 12 non-/s/-stop tokens, the duration of the consonant in the correctly produced cluster ($M = 168.6$ ms, $SD = 36.5$ ms) was significantly shorter than that of the segment in the deletion token ($M = 228.0$ ms, $SD = 46.2$ ms), $t(11) = 3.40, p < .01$. For the three items with /s/ deletion in /s/-stop clusters, the stops were produced with VOTs consistent with voiceless unaspirated consonants (e.g., the /t/ in *steam* → [tim] was 60 ms, compared with 25 ms for /t/ in accurately produced *steam*). We were unable to record P4 producing singleton onsets for comparison. However, given that there was a consistent pattern in which his deletion errors were different from the accurately produced cluster, and in particular because this pattern was produced more similarly to singleton onsets in general, we are confident that these data accurately reflect that the locus of P4's errors arose in phonological processing (see Figure 1d).

Repetition Training Accuracy Changes

To determine the effect of training, we compared pretest and posttest accuracy on the consonant clusters

Figure 1. Duration measures of consonants for each participant produced as singletons, as sounds in clusters, and as sounds in words with consonant deletion from a target cluster. Asterisks indicate a significant difference ($\alpha = .05$) from the duration in the deletion token.

that were tested, separating the trained words from the untrained words. We combined the data for the two pretest and posttest sessions for P2 and P4. Figure 2 presents the accuracy change from pretest to posttest for all participants. P1 and P2 showed sizable increases in accuracy between the pretests and posttests, each increasing their accuracy more than 10% for both trained and untrained items. For P1, this reflected a significant increase for both trained items [pretest: 59.5%, 78/131; posttest: 72.5%, 95/131; $\chi^2(1) = 4.35, p < .05$] and untrained items [pretest: 55.7%, 73/131; posttest: 70.9%,

93/131; $\chi^2(1) = 5.94, p < .05$]. In a similar way, the increase for P2 was significant with trained items [pretest: 47.1%, 177/376; posttest: 55.3%, 208/376; $\chi^2(1) = 4.79, p < .05$] as well as with untrained items [pretest: 41.9%, 72/172; posttest: 53.4%, 92/172; $\chi^2(1) = 4.21, p < .05$].

In contrast, neither P3 nor P4 showed differences as a result of training. For P3, no differences were obtained between the two testing periods for trained items [pretest: 65.6%, 86/131; posttest: 67.9%, 89/131; $\chi^2(1) = 0.69, ns$] or untrained items [pretest: 77.0%, 101/131; posttest: 70.9%, 93/131; $\chi^2(1) = 0.97, ns$]. In a similar way, no significant differences for P4 were found for either trained items [pretest: 81.7%, 307/376; posttest: 83.0%, 312/376; $\chi^2(1) = 2.52, ns$] or untrained items [pretest: 87.2%, 150/172; posttest: 82.6%, 142/172; $\chi^2(1) = 1.11, ns$].

Figure 2. Change in cluster production accuracy for each individual in the study from pretest to posttest.

General Discussion

The present study systematically investigated whether the criteria for distinguishing phonological and motor deletion errors laid out in our previous work (Buchwald & Miozzo, 2011, 2012) can also predict whether individuals improve their cluster production following a repetition-based training protocol that incorporates principles of motor learning associated with the structure of practice (Maas et al., 2008). Four individuals with aphasia and AOS participated in a repetition training procedure involving eight training sessions, a pretest, and a posttest. Participant

productions were used both to identify the predominant type of errors they produced on the basis of our previous criteria and to evaluate changes in performance. The results revealed a clear and consistent pattern across all four participants. The two individuals whose errors fit the pattern associated with motoric locus of error (i.e., their deletion errors in clusters retained the timing of a cluster) showed significant improvement as a result of the repetition training, being more successful at producing both trained items and untrained items matched for consonant cluster. In contrast, the two individuals whose errors were consistent with errors arising in phonological processing (i.e., their deletion errors in clusters were produced with the timing of singleton consonants) exhibited no change in performance as a result of the repetition training.

These results provide crucial additional support for our previous findings (Buchwald & Miozzo, 2011, 2012) that the acoustic properties in productions with deletion errors can indicate the predominant type of error that a participant makes, and that this can predict the extent to which individuals may benefit from a very basic intervention grounded in principles of motor learning. In the remainder of the discussion, we consider the implications of this work for clinical and theoretical purposes, given the strengths and limitations of the present study.

Implications for Accounts of Spoken Production Disorders

Our approach in this line of research is to combine neurocognitive theories of spoken production with experimental investigation of spoken production disorders. Most neurocognitive accounts of spoken production include distinct but interacting systems corresponding to the distinction between phonological and motoric processing discussed in this article (see Goldrick, Ferreira, & Miozzo, 2014, and articles within) and discussed further in the next section. The existence of these distinct levels makes the prediction that we should be able to find individuals with deficits affecting computations at each level. Our first question was whether we are able to detect these differences in the properties of errors that individuals make, using a broader set of errors than in our original investigations (Buchwald & Miozzo, 2011, 2012). We found a clear distinction between speakers whose remaining consonant in deletion errors matched the timing of that sound in the cluster (P1 and P2) and speakers whose errors were shorter than the timing in clusters (P3 and P4). This finding replicates and extends our previous work by adding additional participants and a wider range of deletion error contexts.

Although the investigation of differences in error acoustics is based in neurocognitive theories, the clinical significance of this distinction is indicated by the clear and consistent findings that repetition training leads to improved performance only for individuals with predominantly motor errors. This finding suggests that the repetition training, when consistent with the structure of practice that best promotes motor learning (Maas et al., 2008), provides some

rehabilitation for the motor system, as indicated by a reduced number of deletion errors in two individuals with AOS and concomitant aphasia posttraining (P1 and P2). In addition, to the degree that the repeated production of clusters is related to other articulatory approaches to AOS treatment, this finding is also consistent with a large body of evidence on sound production treatment, which indicates that articulatory–kinematic approaches to AOS are effective because these approaches provide rehabilitation for the system that generates these motor errors (Ballard et al., 2015; Wambaugh et al., 2016). Taken together, the acoustic and repetition training findings indicate that our approach to determining error types is meaningful from both neurocognitive and clinical perspectives.

It is worth noting that our findings do not simply restate the clinical distinctions between AOS and phonological impairment in aphasia from the literature. Similar to our previous results (Buchwald & Miozzo, 2012), we found that at least one of the speakers in our study (P3)—who met the diagnostic criteria for AOS—produced errors that appeared to have a phonological locus according to our account outlined in the Introduction, at least with regard to deletion errors. Furthermore, this individual did not benefit from repeated exposure to the clusters in a repetition task. Thus, there are meaningful distinctions that can be identified even among individuals with complex disorders that include components of AOS and aphasia. Although our work has covered a relatively small number of participants to date, it suggests that the population of those who have AOS with aphasia may include individuals who make sound deletion errors that arise in phonological processing, and that the presence of these errors may be related to whether these individuals benefit from a basic repetition training. One approach to AOS research has been to compare individuals with aphasia with AOS to individuals with aphasia without AOS on tasks that tap into the motor system (Haley et al., 2013; Maas et al., 2014, 2015). In some cases, the data obtained in these studies are equivocal, with only some members of the AOS with aphasia population exhibiting the expected behavior (e.g., Maas et al., 2015). Thus, for research that involves individuals with co-occurring deficits, it may be important to provide additional testing of the errors of the individuals with AOS and aphasia in order to ensure the internal consistency of the populations. In other words, although this is a population noted for its variability, it remains critical to explore how much of that variability is structured to the extent that we can detect reliable and clinically significant distinctions.

Interactivity in Spoken Production

Another open question raised by this work relates to interactivity among spoken production processes. As mentioned in the Introduction, there is clear and consistent evidence that some spoken production errors reflect properties of the target and the error, which suggests that the activation of the target influences the error response through cascading activation (unimpaired participants: Goldrick &

Blumstein, 2006; McMillan & Corley, 2010; Goldrick, Baker, Murphy, & Baese-Berk, 2011; Pouplier, Marin, & Waltl, 2014; impaired participants: Kurowski & Blumstein, 2016; Laganaro, 2012). For example, Goldrick and Blumstein (2006) observed that voicing errors (e.g., /k/ → [g]) were produced with longer VOTs than accurately produced voiced consonants (e.g., /g/ → [g]), revealing the partial activation from the target /k/ in the case of the errors, and a similar finding was reported in the errors of individuals with aphasia (Kurowski & Blumstein, 2016). It has been argued that cascading activation occurs when the activation at one level of processing affects processing at a later level, which typically refers to whether an unproduced item is active (Rapp & Goldrick, 2000) but may also relate to whether the activation strength at one level cascades down to a later level (Buchwald & Falconer, 2014). Thus, cascading activation requires the existence of two distinct levels that interact with one another (Rapp & Goldrick, 2000).

As discussed in the Introduction, we are convinced that there is interactivity in the speech production system and do not intend to endorse a strictly modular account of spoken production (as discussed in the Introduction). In our view, the evidence supports an interactive system that contains interactivity but is “staged” in the sense that Dell et al. (1997) describe (also see Nozari et al., 2010). In these frameworks, there is simultaneous activity of elements at different levels (e.g., word level, phoneme level), but once a node is selected for production at a higher level (e.g., word level), that node receives a jolt of activation that spreads to its connected nodes at lower levels (e.g., phoneme level). Thus, although nodes corresponding to competitors at lower levels may retain some activation, the nodes associated with the selected item receive additional activation that in turn allows them to be selected accurately despite the interactivity.

We note that there are other accounts of lexical knowledge based on connectionist spreading-activation frameworks in which parallel activation occurs throughout lexical processing (e.g., Kendall, Oelke, Brookshire, & Nadeau, 2015), and thus there may not be a staged component to these accounts. A full evaluation of these proposals and how they relate to the present data or the other lexical access data addressed by the frameworks discussed above is outside of the scope of this article.

Given the evidence for interactivity, it is not clear why the phonological errors of the individuals we have observed do not appear to show any trace of the target in their responses. In particular, we would expect that the errors arising in phonology would be somewhat distinct from a target singleton but still distinguishable from the timing of a cluster. We do not have a strong explanation of the lack of this as a clear pattern in the data. Perhaps the most salient difference between our data and the data supporting cascading activation is that those productions have primarily involved substitution errors, whereas our data focus on deletions given their frequency and the straightforward nature of the predictions for these errors. Thus, it is possible that interactivity affects different error types

differently. In addition, given the fact that these speakers are more than a year past the CVA onset, it is possible that the data we report on are patterns that have developed over time on the basis of stable production dynamics. On this point, we would simply note that these deletion errors remain relatively rare and are thus unlikely to be the result of an explicit strategy. However, even though these errors are rare, the acoustic properties have predicted responsiveness to a repetition-based intervention. Although the data we have obtained have been clear and consistent, this issue will need to be investigated further.

Limitations and Future Directions

This work replicated and extended our previous work by using additional participants, a wider range of clusters, and a more systematic approach to determining whether the participants who make motor errors are improving following the repetition-based intervention task. However, there are some clear limitations to the approach we used here that need to be addressed in future work. First and foremost, the design of future work should follow the best practices for single-subject intervention research, including multiple baselines, multiple probes, and establishing control within and across speakers. Although the current study was intended to provide further grounding for that work, our baseline and posttest measures were limited, and our midtreatment measures came from the practice itself and not from probes. Furthermore, we did not establish experimental control across or within participants, which is critical to ensure that the improvement is the direct result of the repetition-based intervention.

Even within the type of study we performed, there were some limitations, including the differences in how participants were tested and the lack of a follow-up test. In addition, although the repetition sessions when singletons were recorded were conducted in the same manner as the primary training and testing, it would be preferable for these words to have been recorded as part of the same testing session to eliminate the possibility of differences arising on the basis of the dates of testing. Furthermore, although this work has suggested that the acoustic properties of deletion errors can help to predict who will respond favorably to a repetition-based intervention, these were a small subset of the errors made by each individual, and it remains possible that they differ from the other errors made by each speaker.

Although the approach here consisted of running the participants on the tasks and using their performance to identify their predominant error type, it is noteworthy that there was a difference in the aphasia profiles of these individuals. In particular, P1 and P2 presented with nonfluent aphasia whereas P3 and P4 presented with fluent aphasia. The degree to which this directly relates to our findings is unclear. In our previous work (Buchwald & Miozzo, 2011, 2012), we reported on one individual with nonfluent aphasia whose errors fit the phonological pattern, ruling out the possibility that the patterns we reported here are based solely on the fluent versus nonfluent distinction. It thus

remains critical to test even more speakers to determine the relationship between the pattern of deletion errors and their response to the repetition-based intervention.

Conclusion

The work presented here provides crucial support for claims that we can identify the source of sound production errors from the details of the errors themselves using a theory-driven approach. Furthermore, we provide additional support that the distinction between the error types is a clinically meaningful one, predicting improvement on the basis of a repetition training protocol. These promising results still must be considered against the backdrop of the complexity of sound production deficits, and it remains possible that not all individuals will consistently produce the same type of error. As we move forward in this work, we hope to test the limits of this approach and to determine whether some individuals produce both error types. For these speakers, we would expect the motor errors, but not the phonological errors, to decrease following repetition training. Future directions also include finding additional approaches to treating the phonological errors that are seen in these speakers.

Acknowledgments

The research and preparation of this article were supported by a grant from the NYU University Research Challenge Fund and National Institutes of Health Grant K01DC014298 to Adam Buchwald. The authors would like to thank Rebecca Martinez, Toby Shaw, Jennifer Wang, Karen Hong, and Brittany Sugzda for help with data collection and analysis.

References

- Aichert, I., & Ziegler, W. (2004). Syllable frequency and syllable structure in apraxia of speech. *Brain and Language*, 88, 148–159.
- 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, 316–337. https://doi.org/10.1044/2015_AJSLP-14-0118
- Bislick, L. P., Weir, P. C., Spencer, K., Kendall, D., & Yorkston, K. M. (2012). Do principles of motor learning enhance retention and transfer of speech skills? A systematic review. *Aphasiology*, 26, 709–728. <https://doi.org/10.1080/02687038.2012.676888>
- Boersma, P., & Weenink, D. (2015). Praat: Doing phonetics by computer (Version 6.0.03) [Computer program]. Retrieved from <http://www.praat.org/>
- Browman, C. P., & Goldstein, L. M. (1986). Towards an articulatory phonology. *Phonology*, 3, 219–252.
- Browman, C. P., & Goldstein, L. M. (1990). Articulatory gestures as phonological units. *Phonology*, 6, 201–251.
- Buchwald, A., & Falconer, C. (2014). Cascading activation from lexical processing to letter-level processing in written word production. *Cognitive Neuropsychology*, 31, 606–621. <https://doi.org/10.1080/02643294.2014.951609>
- Buchwald, A., & Miozzo, M. (2011). Finding abstract structure in word production: Evidence from aphasia. *Psychological Science*, 22, 1113–1119.
- Buchwald, A., & Miozzo, M. (2012). Phonological and motor errors in individuals with acquired impairment. *Journal of Speech, Language, and Hearing Research*, 55, 1573–1586.
- Buchwald, A., Rapp, B., & Stone, M. (2007). Insertion of discrete phonological units: An ultrasound investigation of aphasic speech. *Language and Cognitive Processes*, 22(6), 910–948.
- Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. New York, NY: Harper and Row.
- Darley, F. L., Aronson, A. E., & Brown, J. R. (1975). *Motor speech disorders*. Philadelphia, PA: Saunders.
- Dell, G. (1986). A spreading activation theory of retrieval in sentence processing. *Psychological Review*, 93, 283–321.
- Dell, G. (1988). The retrieval of phonological forms in production: Tests of predictions from a connectionist model. *Journal of Memory and Language*, 27, 124–142.
- Dell, G., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review*, 104, 801–838.
- Duffy, J. R. (2005). *Motor speech disorders: Substrates, differential diagnosis, and management*. St. Louis, MO: Elsevier Mosby.
- Falconer, C., & Buchwald, A. (2013). Does orthographic overlap influence lexical selection? *Aphasiology*, 27, 849–866.
- Fodor, J. (1983). *Modularity of mind: An essay on faculty psychology*. Cambridge, MA: MIT Press.
- Galluzzi, C., Bureca, I., Guariglia, C., & Romani, C. (2015). Phonological simplifications, apraxia of speech and the interaction between phonological and phonetic processing. *Neuropsychologia*, 71, 64–83. <https://doi.org/10.1016/j.neuropsychologia.2015.03.007>
- Goldrick, M., Baker, H. R., Murphy, A., & Baese-Berk, M. (2011). Interaction and representational integration: Evidence from speech errors. *Cognition*, 121, 58–72.
- 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.
- Goldrick, M., Ferreira, V., & Miozzo, M. (2014). *The Oxford handbook of speech production*. Oxford, UK: Oxford University Press.
- Goldrick, M., & Rapp, B. (2007). Lexical and post-lexical phonological representations in spoken production. *Cognition*, 102, 219–260.
- Goldstein, L., Byrd, D., & Saltzman, E. (2006). The role of the vocal tract gestural action units in understanding the evolution of phonology. In M. A. Arbib (Ed.), *Action to language via the mirror neuron system* (pp. 215–249). Cambridge, UK: Cambridge University Press.
- Goodglass, H., Kaplan, E., & Barresi, B. (2001). *The Assessment of Aphasia and Related Disorders* (3rd ed.). Philadelphia, PA: Lippincott Williams & Wilkins.
- Guenther, F. H., & Perkell, J. S. (2004). A neural model of speech production and its application to studies of the role of auditory feedback in speech. In B. Maassen, R. Kent, H. Peters, P. van Lieshout, & W. Hulstijn (Eds.), *Speech motor control in normal and disordered speech* (pp. 29–49). Oxford UK: Oxford University Press.
- 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.

- Jakobson, R. (1968). *Child language, aphasia, and phonological universals*. The Hague, the Netherlands: Mouton. (Original work published 1941)
- Kaplan, E., Goodglass, H., & Weintraub, S. (1983). *Boston Naming Test*. Philadelphia, PA: Lea & Febiger.
- Kay, J., Lesser, R., & Coltheart, M. (1992). *PALPA: Psycholinguistic Assessments of Language Processing in Aphasia*. Hove, England: Lawrence Erlbaum Associates.
- Kendall, D. L., Oelke, M., Brookshire, C. E., & Nadeau, S. E. (2015). The influence of phonomotor treatment on word retrieval abilities in 26 individuals with chronic aphasia: An open trial. *Journal of Speech, Language, and Hearing Research*, 58, 798–812. https://doi.org/10.1044/2015_JSLHR-L-14-0131
- Kertesz, A. (2006). *Western Aphasia Examination-Revised*. San Antonio, TX: Pearson.
- Kurowski, K., & Blumstein, S. E. (2016). Phonetic basis of phonemic paraphasias in aphasia: Evidence for cascading activation. *Cortex*, 75, 193–203. <https://doi.org/10.1016/j.cortex.2015.12.005>
- Laganaro, M. (2005). Syllable frequency effect in speech production: Evidence from aphasia. *Journal of Neurolinguistics*, 18, 221–235.
- Laganaro, M. (2008). Is there a syllable frequency effect in aphasia or in apraxia of speech or both?. *Aphasiology*, 22(11), 1191–1200.
- Laganaro, M. (2012). Patterns of impairments in AOS and mechanisms of interaction between phonological and phonetic encoding. *Journal of Speech, Language, and Hearing Research*, 55(Suppl.), S1535–S1543. [https://doi.org/10.1044/1092-4388\(2012/11-0316\)](https://doi.org/10.1044/1092-4388(2012/11-0316))
- Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22, 1–75.
- Maas, E., Gutiérrez, K., & Ballard, K. J. (2014). Phonological encoding in apraxia of speech and aphasia. *Aphasiology*, 28, 25–48. <https://doi.org/10.1080/02687038.2013.850651>
- 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*, 58, 185–200. https://doi.org/10.1044/2014_JSLHR-S-13-0300
- Maas, E., Robin, D. A., Hula, S. N. A., Freedman, S. E., Wulf, G., Ballard, K. J., & Schmidt, R. A. (2008). Principles of motor learning in treatment of motor speech disorders. *American Journal of Speech-Language Pathology*, 17, 277–298.
- McCloskey, M., Macaruso, P., & Rapp, B. (2006). Grapheme-to-lexeme feedback in the spelling system: Evidence from a dysgraphic patient. *Cognitive Neuropsychology*, 23, 278–307.
- McMillan, C. T., & Corley, M. (2010). Cascading influences on the production of speech: Evidence from articulation. *Cognition*, 117(3), 243–260.
- McNeil, M. R., Odell, K., Miller, S. B., & Hunter, L. (1995). Consistency, variability, and target approximation for successive speech repetitions among apraxic, conduction aphasic, and ataxic dysarthria speakers. *Clinical Aphasiology*, 23, 39–55.
- McNeil, M. R., Pratt, S. R., & Fossett, T. R. D. (2004). The differential diagnosis of apraxia of speech. In B. Maassen, R. Kent, H. Peters, P. van Lieshout, & W. Hulstijn (Eds.), *Speech motor control in normal and disordered speech* (pp. 389–414). Oxford, UK: Oxford University Press.
- 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, NY: Thieme.
- Miozzo, M., & Buchwald, A. (2013). On the nature of sonority in spoken word production: Evidence from neuropsychology. *Cognition*, 128, 287–301.
- Miozzo, M., Pulvermüller, F., & Hauk, O. (2015). Early parallel activation of semantics and phonology in picture naming: Evidence from a multiple linear regression MEG study. *Cerebral Cortex*, 25, 3343–3355. <https://doi.org/10.1093/cercor/bhu137>
- Morsella, E., & Miozzo, M. (2002). Evidence for a cascade model of lexical access in speech production. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 28, 555–563.
- Navarrete, E., Peressotti, F., Lerose, L., & Miozzo, M. (2017). Activation cascading in sign production. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 43, 302–318.
- Nozari, N., Dell, G. S., & Schwartz, M. F. (2011). Is comprehension necessary for error detection? A conflict-based account of monitoring in speech production. *Cognitive Psychology*, 63, 1–33. <https://doi.org/10.1016/j.cogpsych.2011.05.001>
- Nozari, N., Kittredge, A. K., Dell, G. S., & Schwartz, M. F. (2010). Naming and repetition in aphasia: Steps, routes, and frequency effects. *Journal of Memory and Language*, 63, 541–559. <https://doi.org/10.1016/j.jml.2010.08.001>
- 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, 345–359.
- 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, 67–80.
- O'Shaughnessy, D. (1974). Consonant durations in clusters. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 22, 282–295.
- Pouplier, M., Marin, S., & Waihl, S. (2014). Voice onset time in consonant cluster errors: Can phonetic accommodation differentiate cognitive from motor errors? *Journal of Speech, Language, and Hearing Research*, 57, 1577–1588.
- Prince, A., & Smolensky, P. (2004). *Optimality theory: Constraint interaction in generative grammar*. (Original work published 1993). Retrieved from <http://roa.rutgers.edu/files/537-0802/537-0802-PRINCE-0-0.PDF>
- Rapp, B., & Goldrick, M. (2000). Discreteness and interactivity in spoken word production. *Psychological Review*, 107, 460–499.
- Roach, A., Schwartz, M. F., Martin, N., Grewal, R. S., & Brecher, A. (1996). The Philadelphia Naming Test: Scoring and rationale. *Clinical Aphasiology*, 24, 121–133.
- Romani, C., Galluzzi, C., Bureca, I., & Olson, A. (2011). Effects of syllable structure in aphasic errors: Implications for a new model of speech production. *Cognitive Psychology*, 62, 151–192.
- Schmidt, R. A., & Lee, T. (2005). *Motor control and learning: A behavioral emphasis*. Champaign, IL: Human Kinetics.
- 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(Suppl.), S1544–S1561. [https://doi.org/10.1044/1092-4388\(2012/11-0319\)](https://doi.org/10.1044/1092-4388(2012/11-0319))
- 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.
- Stemberger, J. P. (1985). An interactive activation model of language production. In A. W. Ellis (Ed.), *Progress in the psychology of language* (Vol. 1, pp. 143–141, 186). Hillsdale, NJ: Erlbaum.
- van der Merwe, A. (2009). A theoretical framework for the characterization of pathological speech motor control. In M. R. McNeil

- (Ed.), *Clinical management of sensorimotor speech disorders* (2nd ed., pp. 3–18). New York, NY: Thieme.
- Varley, R. A., & Whiteside, S. P. (2001). What is the underlying impairment in acquired apraxia of speech? *Aphasiology*, 15, 39–49.
- Wambaugh, J. L., Duffy, J. R., McNeil, M. R., Robin, D. A., & Rogers, M. A. (2006a). Treatment guidelines for acquired apraxia of speech: A synthesis and evaluation of the evidence. *Journal of Medical Speech-Language Pathology*, 14, 15–32.
- Wambaugh, J. L., Duffy, J. R., McNeil, M. R., Robin, D. A., & Rogers, M. A. (2006b). Treatment guidelines for acquired apraxia of speech: Treatment descriptions and recommendations. *Journal of Medical Speech-Language Pathology*, 14, 33–67.
- Wambaugh, J. L., Mauszycki, S., & Ballard, K. (2013). Advances in the treatment for acquired apraxia of speech. *Perspectives on Neurophysiology and Neurogenic Speech and Language Disorders*, 23, 112–119. <https://doi.org/10.1044/nnsld23.3.112>
- Wambaugh, J. L., Nessler, C., Wright, S., & Mauszycki, S. C. (2014). Sound production treatment: Effects of blocked and random practice. *American Journal of Speech-Language Pathology*, 23(Suppl.), S225–S245.
- Wambaugh, J. L., Nessler, C., Wright, S., Mauszycki, S., & DeLong, C. (2016). Sound production treatment for acquired apraxia of speech: Effects of blocked and random practice on multisyllabic word production. *International Journal of Speech-Language Pathology*, 18, 450–464. <https://doi.org/10.3109/17549507.2015.1101161>
- Wertz, R. T., Lapointe, L. L., & Rosenbek, J. C. (1984). *Apraxia of speech in adults: The disorder and its management*. Orlando, FL: Grune and Stratton.
- Ziegler, W., & Aichert, I. (2015). How much is a word? Predicting ease of articulation planning from apraxic speech error patterns. *Cortex*, 69, 24–39.
- Ziegler, W., Staiger, A., & Aichert, I. (2010). Apraxia of speech: What the deconstruction of phonetic plans tells us about the construction of articulate language. In B. Maassen & P. van Lieshout (Eds.), *Speech motor control* (pp. 3–22). Oxford, UK: Oxford University Press.