Research Article

Feedforward and Feedback Control in Apraxia of Speech: Effects of Noise Masking on Vowel Production

Edwin Maas,[a] Marja-Liisa Mailend,[a] and Frank H. Guenther[b]

Purpose: This study was designed to test two hypotheses about apraxia of speech (AOS) derived from the Directions Into Velocities of Articulators (DIVA) model (Guenther et al., 2006): the feedforward system deficit hypothesis and the feedback system deficit hypothesis.
Method: The authors used noise masking to minimize auditory feedback during speech. Six speakers with AOS and aphasia, 4 with aphasia without AOS, and 2 groups of speakers without impairment (younger and older adults) participated. Acoustic measures of vowel contrast, variability, and duration were analyzed.
Results: Younger, but not older, speakers without impairment showed significantly reduced vowel contrast with noise masking. Relative to older controls, the AOS group showed

longer vowel durations overall (regardless of masking condition) and a greater reduction in vowel contrast under masking conditions. There were no significant differences in variability. Three of the 6 speakers with AOS demonstrated the group pattern. Speakers with aphasia without AOS did not differ from controls in contrast, duration, or variability.
Conclusion: The greater reduction in vowel contrast with masking noise for the AOS group is consistent with the feedforward system deficit hypothesis but not with the feedback system deficit hypothesis; however, effects were small and not present in all individual speakers with AOS. Theoretical implications and alternative interpretations of these findings are discussed.

A praxia of speech (AOS) is a neurogenic motor speech disorder characterized by slow speech rate, speech sound distortions, sound and syllable segmentation, inconsistent presence but relatively consistent type and location of errors, and dysprosody (Duffy, 2005; McNeil, Robin, & Schmidt, 2009). A consensus exists that AOS reflects an impairment of planning and/or programming speech movements (Deger & Ziegler, 2002; Duffy, 2005; Maas, Robin, Wright, & Ballard, 2008; Van der Merwe, 2009). However, the precise nature of the disorder remains poorly understood, in part because models of speech motor planning have often been underspecified (Ziegler, 2002). The present study was designed to advance our understanding of AOS in the context of a well-developed theory of speech motor control (e.g., Perkell, 2012) and a corresponding computational model that implements aspects of this theory (e.g., Guenther, Ghosh, & Tourville, 2006). This

framework provides a strong foundation for developing and testing hypotheses about impaired speech motor control in AOS and its neural underpinnings.

Speech Motor Control

Recent years have seen significant progress in our understanding of speech motor control. One relatively comprehensive and well-supported framework (Guenther et al., 2006; Perkell, 2012) makes several key assumptions, two of which are most relevant to the present study. First, targets for speech motor control are regions in auditory and somatosensory space, with auditory space considered the primary planning space (Guenther, Hampson, & Johnson, 1998; Perkell, 2012). Target regions specify the range of acceptable variation along relevant dimensions for a given speech sound (e.g., acceptable range of formant values for vowel /i/). Second, speech motor control combines feedback control and feedforward control (Guenther et al., 2006; Perkell, 2012). The Directions Into Velocities of Articulators (DIVA) computational model implements these assumptions for control of speech segments (e.g., Guenther et al., 2006).

[a]University of Arizona, Tucson
[b]Boston University, MA

Correspondence to Edwin Maas: emaas@email.arizona.edu

In the DIVA model, production of a speech sound begins with activation of a speech sound map (SSM) cell in left inferior frontal cortex. The SSM cells represent speech sounds (which may range in size from phonemes to syllables to frequent words and phrases) and are presumed to be activated by higher level input from the phonological encoding stage (Bohland, Bullock, & Guenther, 2010; Guenther et al., 2006). The activated SSM cell then activates a feedforward control system and a feedback control system whose motor commands are combined in the primary motor cortex. Feedback control involves comparing actual auditory and somatosensory feedback signals with expected auditory and somatosensory consequences and generating corrective motor commands to the motor cortex when a mismatch (error) is detected. Expected sensory consequences are encoded as regions in auditory space (superior temporal gyrus) and somatosensory space (postcentral and supramarginal gyri). Feedforward control involves predictive motor commands from the SSM to the motor cortex. Feedforward commands are learned by incorporating the feedback system's corrective commands from previous productions. With sufficient practice, the feedforward commands generate few to no errors so that contributions of the feedback control system are minimal during normal speech, although feedback may be continuously monitored for deviations from expectations, even in adult speakers (Tourville, Reilly, & Guenther, 2008).

Support for the role of feedforward motor control comes from observations that speakers maintain segmental contrasts even when auditory feedback is no longer available (e.g., in individuals who were postlingually deafened or in individuals with unimpaired hearing under conditions of auditory feedback masking; e.g., Lane et al., 2005) and from findings that speakers reveal predictable and persistent changes in response to gradual and systematic perturbations of their auditory feedback (e.g., Houde & Jordan, 1998; Shiller, Sato, Gracco, & Baum, 2009; Villacorta, Perkell, & Guenther, 2007). The fact that speakers continue to produce altered speech for some time even after feedback is restored to normal (so-called aftereffects) indicates that the predictive motor commands have been altered (e.g., Villacorta et al., 2007). Support for the role of feedback control in speech production comes from observations that speech quality may deteriorate to some extent in the absence or attenuation of feedback (e.g., individuals who were postlingually deafened, noise masking; Lane et al., 2005; Perkell, 2012) and from experimental findings of immediate, within-trial compensations in response to sudden alterations of auditory feedback (e.g., Cai, Ghosh, Guenther, & Perkell, 2011; Purcell & Munhall, 2006; Tourville et al., 2008).

Apraxia of Speech and the DIVA Model

In the framework of the DIVA model, there are several possible loci for the deficit in AOS. In this study, we examined two different hypotheses, namely the feedforward system deficit hypothesis (FF hypothesis) and the feedback system deficit hypothesis (FB hypothesis). According to the FF hypothesis, AOS reflects a disruption of feedforward control, whereas feedback control is spared and plays a more prominent role in achieving and maintaining segmental contrast in AOS (Jacks, 2008; cf. also Rogers, Eyraud, Strand, & Storkel, 1996). This hypothesis is consistent with primary features of AOS such as a slow rate and sound distortions (Jacks, 2008) and with the putative left frontal cortical lesions (e.g., Hillis et al., 2004). If the feedforward mechanism is impaired, motor commands arising from the feedforward system would produce errors (e.g., distortions, articulatory groping; cf. Bohland et al., 2010), which in turn would increase the contribution of feedback-based corrective commands to the overall motor command. A greater reliance on feedback control could account for the slower rate due to the need to process and incorporate the feedback signals (Rogers et al., 1996) and might lead to increased spatial and temporal variability of articulatory movements due to corrections needed to counter incorrect feedforward commands (Jacks, 2008). Finally, because learning speech sounds involves updating feedforward commands on the basis of corrective feedback commands, impaired feedforward control is also consistent with the common clinical observation that improvements in speech production often require considerable time and effort for individuals with AOS.

In contrast, the FB hypothesis states that AOS involves impaired auditory feedback control. The suggestion that AOS may involve a disruption of feedback processing is not new (e.g., Bartle-Meyer & Murdoch, 2010; Kent & Rosenbek, 1983; Mlcoch & Noll, 1980; Rogers et al., 1996), but few studies have directly tested this hypothesis. It appears clear that AOS is not an impairment of auditory perceptual processing per se because speakers with AOS often recognize their errors even if they cannot correct them (Deal & Darley, 1972; Kent & Rosenbek, 1983). In addition, speakers with pure AOS did not differ from intact speakers on a range of auditory perceptual tasks, unlike speakers with aphasia without AOS (Square, Darley, & Sommers, 1981). Thus, the difficulty in feedback control may rather be in deriving error information from mismatching feedback and/or generating corrective commands on the basis of such errors (cf. Rogers et al., 1996), perhaps because incorrect target regions are activated, because the internal model that governs corrections is damaged, or because feedback commands cannot be integrated with feedforward commands. Depending on the nature of the feedback deficit, difficulties with using feedback may lead to several features often noted in AOS, including articulatory groping, speech sound distortions, and increased variability. It is interesting that Ballard and Robin (2007) showed, in a visuomotor tracking task, that individuals with AOS were more variable and less accurate in tracking a sinusoidal target with their jaw when feedback (in the form of a visual trace on the screen) was present than when it was not, especially for timing accuracy. Although this was not a speech task and involved novel movements, the results do suggest that integrating feedback with ongoing movements may be impaired in AOS.

The Present Study

In this study we tested behavioral predictions of these two hypotheses by using auditory feedback masking, based on the premise that noise masking would effectively eliminate the auditory feedback signal for use in controlling speech movements (Rogers et al., 1996). Although we recognize that complete removal of all feedback is virtually impossible[1] (Kelso & Tuller, 1983; Kent, Kent, Weismer, & Duffy, 2000), it is reasonable to expect a greater contribution of feedforward control in the absence of auditory information, considering the goal of speech movements (to produce an acoustic signal perceivable by the listener; Guenther et al., 1998), especially for sounds with clear auditory feedback such as vowels.

It is well known that auditory feedback masking produces changes in suprasegmental aspects of speech production (e.g., Lane & Tranel, 1971; Van Summers, Pisoni, Bernacki, Pedlow, & Stokes, 1988). More recently, research has also shown changes in segmental aspects of speech under conditions of auditory feedback masking (e.g., Lane et al., 2005, 2007; Perkell et al., 2007; Van Summers et al., 1988). For instance, Perkell et al. (2007) showed that speakers' acoustic segmental contrasts for vowels and fricatives decrease when their acoustic feedback is masked by speech-shaped noise (suggesting a contribution of auditory feedback in unmasked conditions), even though contrasts did not collapse completely (suggesting sufficiently robust feedforward commands to support segmental contrasts). Such experimental demonstrations of maintained but reduced contrast under masking conditions for segmental contrasts are predicted by the DIVA model due to the combined feedback–feedforward control scheme.

Turning to the predictions of our hypotheses, the FF hypothesis predicts that speakers with AOS will show a *disproportionate reduction* of acoustic segmental contrast with auditory masking, relative to unimpaired speakers. The rationale is that masking of auditory feedback, on which speakers with AOS rely to a greater extent than do speakers without AOS, will force a greater reliance on the feedforward system (and somatosensory feedback), thus exposing the feedforward deficit more clearly. In contrast, the FB hypothesis predicts *increased* acoustic contrast under auditory feedback masking or *unchanged* contrast with decreased variability. The rationale is that if apraxic speech characteristics arise due to interference from self-produced auditory feedback, then masking of such auditory feedback should enable the intact feedforward control mechanism to implement the intended speech targets more correctly and/or with greater stability.

To our knowledge, only two published studies have examined speech production under auditory feedback masking conditions in AOS (Deal & Darley, 1972; Rogers et al.,

1996). Deal and Darley (1972) reported no effects of masking on perceptually judged phonemic accuracy. Rogers et al. (1996) used white-noise masking to explore reasons for the well-known finding of longer vowel duration in AOS (e.g., Kent & Rosenbek, 1983). They reasoned that if speakers with AOS prolong vowels in order to verify target attainment using auditory feedback (as in the FF hypothesis), then noise masking should result in shorter vowels because the auditory feedback strategy would be impossible. Two of the three speakers with AOS and two of the three control speakers tended to prolong, not shorten, vowels under masking conditions; this effect was similar in both groups. However, Rogers et al. did not examine segmental contrast and included only three speakers with AOS and three controls. The present study set out to test the predictions of the two hypotheses relative to segmental contrast in six participants with AOS, four with aphasia without AOS, and two groups with unimpaired speech (older and younger adults).

Method

Participants

The study included six participants with AOS and aphasia (AOS group), four participants with aphasia without AOS (APH group), and 12 age-matched (older) control speakers (see Tables 1 and 2 for information about patients and controls, respectively).[2] Eleven younger speakers also completed the study to verify that our procedures were effective by replicating previous findings of contrast reduction with masking. All participants passed a hearing screen at 0.5, 1, 2, and 4 kHz at 40 dB in at least one ear. We adopted this criterion rather than the typical 25-dB criterion (American Speech-Language-Hearing Association, 1997) because the hearing screen was not performed in a sound-treated room and because we wanted to minimize the potential exclusion rate of patients due to age-related hearing loss (Morrell, Gordon-Salant, Pearson, Brant, & Fozard, 1996). All participants had hearing thresholds in the better ear at or below 25 dB at all four frequencies, except one of 12 age-matched controls at 0.5 kHz (30 dB), five of 12 age-matched controls at 4 kHz (range = 30 to 40 dB), and one of four speakers from the APH group at 4 kHz (35 dB). There were no differences between the patient groups and age-matched control group for hearing level in the better ear at any of the frequencies (p values > .05).

Apraxia of speech was diagnosed by an experienced clinical researcher (first author) using a 3-point rating scale (1 = *no AOS*, 2 = *possible AOS*, 3 = *AOS*). Current best practice in diagnosing AOS is expert opinion because there are no operationalized and validated tests or measures available (Wambaugh, Duffy, McNeil, Robin, & Rogers, 2006). To go beyond a single expert's judgment and increase confidence in the diagnosis, two certified speech-language

---

[1]For instance, speakers might also use proprioceptive or tactile information (Ghosh et al., 2010; Honda, Fujino, & Kaburagi, 2002; Leung & Ciocca, 2011) or bone-conduction signal (Reinfeldt, Östli, Håkansson, & Stenfelt, 2010).

[2]The participant codes used here refer to the same patients as those reported in Mailend and Maas (2013). More detailed descriptions of each participant are provided in the online supplemental materials.

Table 1. Patient information.

Characteristic	AOS						Aphasia			
	200	201	203	204	205	206	301	304	306	307
Age	58	68	67	56	59	72	68	64	40	57
Sex	M	M	M	F	F	M	M	F	M	F
Hand	R	R	R	R	R	R	R	R	R	R
Education (years)	12	22	12	12	13	16	15	15	15	12
Language	AE	AE	SE	SE	AE	AE	BE	AE	AE	AE
Hearing[a]	pass	pass	pass	pass	pass	pass	pass	pass	pass	pass
TPO (year;month)	4;6	6;9	2;7	5;6	5;1	7;6	7;10	2;10	3;2	9;11
Etiology	LH CVA	LH CVA	LH CVA	Tumor & LH CVA	LH CVA	LH & RH CVA	LH CVA	LH CVA	LH CVA	LH CVA
Lesion	B; I; PreC	n/a	I; PF; P	FT	n/a	PT; P; IO	B; FT; P	T; O; PF	IFG; T; P	IFG; T; P
Aphasia type[b]	WNL	Anomic	Broca's	Broca's	Anomic	Wernicke	Wernicke	Conduction	Anomic	Conduction
WAB-R AQ[b]	94.2	93.2	50.3	58.7	82.1	69.3	74.9	2-3/5[2]	92	86.3
AOS severity[c]	mild-mod	mild	mod-severe	mild-mod	mild-mod	mild	none	none	none	none
AOS rating[d]	2.7	3.0	3.0	2.7	2.3	2.0	1.0	1.3	1.0	1.3
Dysarthria[e]	none	mild	mild	none	mild	mild	none	none	none	none
Oral apraxia[c]	mild	mild	mild	mild	none	mod	mild	none	none	none
Limb apraxia[c]	none	mild	mild	mild	none	mod	none	none	none	none

Note.   AOS = apraxia of speech; R = right; AE = American English; SE = Spanish–English bilingual; BE = British English; TPO = time post onset; LH = left hemisphere; CVA = cerebrovascular accident; RH = right hemisphere; B = Broca's area; I = insula; PreC = precentral gyrus; n/a = not available; PF= posterior frontal lobe; FT = fronto-temporal; PT = posterior & middle temporal lobe; IO = inferior occipital lobe; T = temporal lobe; O = occipital lobe; IFG = inferior frontal gyrus; P = parietal lobe; WNL = within normal limits; WAB-R = Western Aphasia Battery–Revised; AQ = aphasia quotient; mod = moderate.

[a]Pure-tone hearing screening at 500, 1000, 2000, and 4000 Hz; pass at 40 dB level for better ear. [b]Based on the WAB-R (Kertesz, 2006), except for APH 304 (based on Boston Diagnostic Aphasia Examination–Third Edition; Goodglass et al., 2000). [c]Based on Apraxia Battery for Adults–Second Edition (Dabul, 2000). [d]Mean rating across three diagnosticians (1 = no AOS, 2 = possible AOS, 3 = AOS). [e]Dysarthrias were diagnosed perceptually based on a motor speech exam (Duffy, 2005) and were all of the unilateral upper motor neuron type.

Table 2. Participant information for older and younger controls.

Characteristic	Older controls	Younger controls
N	12	11
Age	66 (5)	22 (5)
Sex	6 F, 6 M	11 F
Hand	10 R, 1 L, 1 A	11 R
Education (years)	18 (3)	15 (2)
Language	11 AE; 1 SE	11 AE

Note. R = right; L = left; A = ambidextrous; AE = American English; SE = Spanish–English bilingual. For age and education, values are means with standard deviations in parentheses.

pathologists experienced with diagnosing motor speech disorders also independently rated audio and video samples of the participants using the same scale and criteria. Diagnosis of AOS was based on the criteria proposed by Wambaugh et al. (2006), judged from speech samples obtained from various speaking tasks, including those from the Apraxia Battery for Adults–Second Edition (Dabul, 2000), the Western Aphasia Battery–Revised (Kertesz, 2006), and conversational speech. In particular, AOS was diagnosed on the basis of presence of the following features: slow speech with longer segment and intersegment durations, dysprosody, distortions and distorted substitutions, and segmental errors that were relatively consistent in type (distortions) and location within the utterance. Normal rate and normal prosody were considered exclusionary criteria for the diagnosis of AOS (Wambaugh et al., 2006). All participants with AOS displayed these primary characteristics; all also showed nondiscriminative behaviors such as articulatory groping, occasional islands of fluent speech, initiation difficulties, more errors on longer words, and self-correction attempts, but these features were not used to diagnose AOS.

Consistent with previous reports (Haley, Jacks, de Riesthal, Abou-Khalil, & Roth, 2012), reliability across diagnosticians was not perfect: Unanimous agreement across all three raters was reached for only four of 10 patients (40%), with 50%, 70%, and 50% unanimous agreement for the three pairs. Agreement within 1 point of the scale was acceptable at 80% (eight of 10) across all three raters (80%, 90%, and 100% agreement for the three pairs of raters). These findings further underscore the need for improved and operationalized criteria in our field (Haley et al., 2012). For purposes of the present study, we operationally classified participants into the AOS group if their mean rating across the three judges was ≥2 (possible or definite AOS; cf. Mailend & Maas, 2013). Mean ratings are provided in Table 1.

Aphasia was diagnosed on the basis of performance on a standardized aphasia test (Western Aphasia Battery–Revised, Kertesz, 2006; or Boston Diagnostic Aphasia Examination–Third Edition, Goodglass, Kaplan, & Barresi, 2000); dysarthria was determined on the basis of a motor speech exam and oral mechanism exam (Duffy, 2005). All participants were native speakers of English (all monolingual except two participants with AOS and one control speaker).

Patients were at least 1 year postonset and had a variety of left hemisphere lesions.

Eight additional participants were recruited but were excluded due to significant hearing loss (four older controls), history of speech problems (one older control), too few analyzable data points in the experimental task due to errors and/or recording or measurement problems (one speaker with AOS), or inability to complete the task due to significant word-finding problems (two speakers with aphasia). All procedures were approved by the University of Arizona Institutional Review Board, and all participants provided informed consent. Participants were recruited from the university's clinic and local hospitals and were compensated $10 per hour for their participation.

Materials

Speech targets included the six vowels /i/, /ɪ/, /ɛ/, /æ/, /ʌ/, /u/ in /bVt/ words (e.g., *beet*) in the carrier phrase "A (target) again." Including four filler words (*shock, sock, sheet, seat*), there were a total of 10 consonant–vowel–consonant (CVC) words. These targets were selected because they produce salient and relatively prolonged auditory feedback and are therefore susceptible to feedback masking and because these targets are amenable to acoustic analysis (Lane et al., 2005; Perkell et al., 2007). The carrier phrase served to reduce possible initiation difficulties affecting the target words; however, the carrier phrase was kept simple to ensure speakers with aphasia would be able to read and produce it. Colored line drawings of the targets were also included to facilitate retrieval and production of the targets.

Equipment

The experiment was controlled by E-Prime software (Version 2; Psychology Software Tools, Inc., Sharpsburg, PA), run on a Dell Inspiron 530 computer with a 21.5″ LCD screen. The experimenter controlled the experiment via a button box (Serial Response Box; Psychology Software Tools, Inc.). Speech-shaped masking noise (Nilsson, Soli, & Sullivan, 1994) was presented over TDH-39 headphones (Telephonics Corporation, Farmingdale, NY) calibrated by a clinical audiologist at 95 dB SPL (Lane et al., 2005; Perkell et al., 2007). An ExTech SL130 audio level meter (ExTech Instruments Corporation, Nashua, NH; settings: A-scale frequency weighting, slow response, amplitude range of 30 to 80 dB) was used to help participants maintain loudness within their habitual range. Speech responses were recorded via an M-Audio Aries condenser microphone (M-Audio, Cumberland, RI) onto a Marantz CDR-420 CD recorder (Marantz America LLC, Mahwah, NJ) at 44.1 kHz.

Task and Procedure

Participants were asked to produce the target phrases in the context of a modified self-selection paradigm (Maas et al., 2008). Each trial involved the following sequence of events (see Figure 1): An asterisk appeared in the center

Figure 1. Schematic outline of trial events (masking trial depicted). ISI = interstimulus interval.
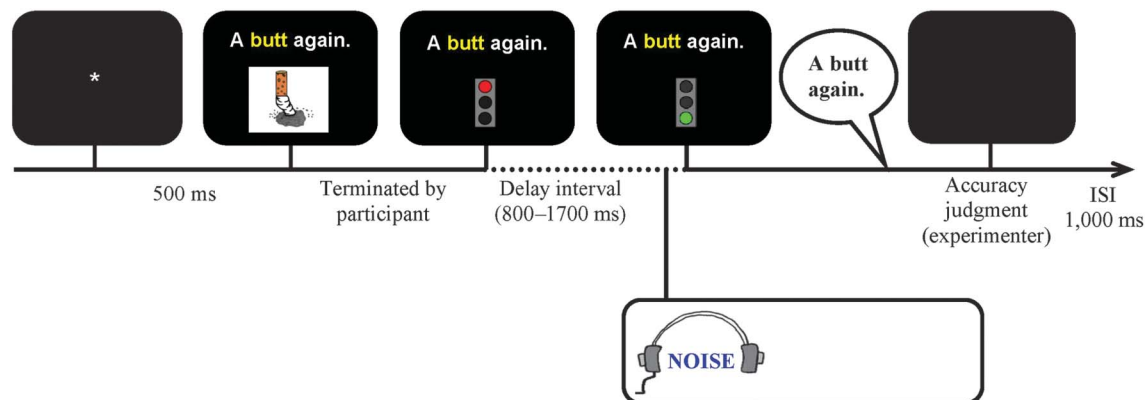
of the screen for 500 ms, followed by the carrier phrase with the target word (in 22-point Arial) above the associated picture. We used both a written carrier phrase and a picture to minimize word-finding difficulties and provide multimodal cues for the target utterance. The phrase and picture remained on the screen until participants pressed the space bar to indicate that they were ready. Next, a delay interval began, during which the phrase remained on the screen and a red traffic light replaced the picture. Participants were to hold their response until the go-signal (green traffic light), which appeared after a delay that randomly varied between 800 and 1,700 ms to prevent initiation before the noise. In normal feedback blocks (*Silence*), participants produced speech with normal feedback without using headphones. In masking blocks (*Noise*), masking noise (with a rise time of 70 ms to minimize risk of hearing damage and annoyance with the sound; Kjellberg, 1990) was presented over headphones from 500 ms before the go-signal throughout the response, terminated by the experimenter's button press to judge response accuracy. Participants 203, 204, 206, and 301 could not reliably produce the carrier phrase and produced the target words in isolation instead. For Participant 204, the experimenter also frequently presented a live auditory model during the preparation interval to facilitate word retrieval. There is no indication that these procedural differences led to different patterns of results.

Incorrect trials were rerun at the end of each block to facilitate collecting equal numbers of usable trials for all targets. Errors were defined as (a) failure to respond, (b) perseveration or production of a different target, (c) semantic paraphasia, (d) unintelligible response, or (e) loudness exceeding the maximum level. Neither distortions nor errors in the phrase were considered errors. Trials on which participants initiated speech during the delay were also rerun. An intertrial interval of 1,000 ms followed the experimenter's judgment of accuracy.

To minimize effects of possible loudness differences, a real-time recording level indicator was visible next to the screen, and speakers were asked to maintain their loudness within the same range in both conditions. For each participant,

a maximum level was specified on the basis of their habitual speaking loudness, determined at the beginning of the experiment. The sound level meter was placed approximately 50 cm from the participant's face. When the maximum loudness level was exceeded, three red lights flashed for 3 s, and the participant was reminded to maintain his or her habitual loudness level (cf. Kelso & Tuller, 1983; Purcell & Munhall, 2006; Rogers et al., 1996). None of the participants demonstrated difficulty with this task requirement.[3]

The experiment consisted of a total of 16 blocks of 10 trials, alternating between normal feedback (Silence) and noise masking blocks (Noise) in order to distribute any potential discomfort from the noise evenly across the experiment (cf. Rogers et al., 1996). All participants started with a Silence block to familiarize them with the task and phrases. Within each block, target phrases were elicited in random order. Between blocks, participants were provided rest intervals of a self-determined duration. Participants were tested individually in a quiet room and were seated comfortably in front of a computer. The experiment took 30 to 60 min.

Design and Analysis
Acoustic Measures

The measures of interest were acoustic measures (described below) on perceptually acceptable tokens. All perceptual errors were excluded from the acoustic analysis, except vowel substitutions, vowel distortions, and consonant voicing errors (this involved 36 trials total, accounting for 6.7% of data for the AOS group, 0.3% for the APH group, 0.3% for older controls, and 0.5% for younger

---

[3]There were only 27 such responses total (0.98%). Seven of these were produced by AOS 200; of the other five participants with AOS, only two made one such error (the others made none). Among the other groups, the range of "too loud" errors was 0 to 3 (younger controls), 0 to 4 (older controls), and 0 to 2 (APH). Analysis including those "too loud" trials for which valid acoustic measures could be obtained did not change any of the findings.

controls).[4] These errors were included in the acoustic analysis because the masking manipulation was hypothesized to affect vowel contrast; thus, excluding vowel errors (including extreme distortions that might be perceived as substitutions) would amount to excluding the data of interest. Furthermore, we did not expect final consonant voicing errors to substantially affect the vowel contrast measures. The pattern of results did not change when these errors were also excluded.

After exclusion of perceptual errors, an additional 123 trials (4.5%) were excluded due to measurement and recording problems (e.g., clipped waveform, no visible formants). Each acceptable vowel token was manually analyzed with Praat (Version 5.0.17; Boersma & Weenink, 2008). Vowel duration was determined from the first to the final vertical striations where at least the first two formants (F1 and F2) were visible on a wideband spectrogram. F1 and F2 were extracted on the basis of the linear predictive coding (LPC) envelopes with a Gaussian-like analysis window of 50 ms and a default bandwidth of 5000 Hz for male speakers and 5500 Hz for female speakers. The peaks from the LPC envelope were verified against the wideband spectrogram. In case of a mismatch between the two measurements, the LPC parameters (e.g., the bandwidth of the analysis window and the maximum number of formants to be searched within the bandwidth) were changed so that the LPC peaks lined up with the most likely formant values visible on the wideband spectrogram. Only those trials where agreement between LPC and spectrographic measures was attained were included in the statistical analysis (1.2% excluded). Next, a MATLAB script (Version 7.11.0; Mathworks, Natick, MA) was used to extract F1 and F2 of each vowel at the time points of 20%, 50%, and 80% of the vowel duration (cf. Jacks, Mathes, & Marquardt, 2010). Finally, the formula $2,595 \times \log10 (1 + (F / 700))$ was used to transform all formant values into mel space (Lane et al., 2007; Perkell et al., 2007).

Intrarater reliability of acoustic measures was assessed for three of 28 participants' complete data (9.4% of data; one younger control, one older control, one participant with AOS), with time between measurements ranging from 3 months to 2 years. The average absolute differences revealed acceptable to high reliability (duration: 9.8 ms; F1: 5.5 Hz; F2: 8.0 Hz), as did the correlations (duration: 0.882; F1: 0.997; F2: 0.998).

Statistical Analysis

Statistical analyses were based only on vowels for which at least four of eight acceptable tokens per condition (Silence, Noise) were available (average number of trials included: 7.5, 7.7, 6.9, 7.4 for younger controls, older controls, AOS, and APH, respectively). This resulted in exclusion

of the vowel /ɪ/ for participant AOS 203. For consistency, all analyses reported herein were based on the remaining five vowels for all participants; the results revealed the same pattern when analyzed with all six vowels excluding AOS 203, indicating that the results were not driven by the vowel /ɪ/.

The primary dependent variable of interest for our hypotheses was average vowel spacing (AVS), a measure of segmental contrast used in previous masking studies (Lane et al., 2005; Perkell et al., 2007). The AVS was calculated for each participant as the average of Euclidian distances in M1 × M2 space between each of the 10 possible vowel pairs (based on each vowel's mean). As such, AVS represents the average intervowel distance, which captures the degree of contrast within the vowel space (see Figure 2). A higher AVS indicates a greater separation between vowels and thus greater vowel contrast. Secondary dependent measures included vowel duration and vowel dispersion. Vowel dispersion represents the token-to-token variability of a vowel around its mean location in M1 × M2 space (Perkell et al., 2007) and was calculated as the average of the Euclidian distances between each vowel token and that vowel's mean.

Given that the younger and older groups with unimpaired speech differed in terms of sex distribution (11 female speakers in the younger group; six female and six male speakers in the older group), data from younger speakers were analyzed separately, using 2 (Condition) × 3 (Time) repeated measures analyses of variance (ANOVAs). Patient data were compared with the older control group both at the group level using ANOVAs and at the individual level using single-case comparison methods (Crawford, Garthwaite, & Porter, 2010). The AVS was analyzed using separate 3 (Group) × 2 (Condition) × 3 (Time) ANOVAs; vowel duration with 3 (Group) × 2 (Condition) × 5 (Vowel) ANOVAs; and vowel dispersion with 3 (Group) × 2 (Condition) × 5 (Vowel) × 3 (Time) ANOVAs. Tukey post hoc tests were used to identify significant effects. In addition, because our hypotheses made predictions specifically about the Group × Condition interaction for the AOS group for AVS, we conducted separate planned ANOVAs for only the AOS group and age-matched controls to address these critical predictions. Individual patients were also compared with the older control group using Bayesian methods for testing differences in each condition (using SingleBayes_ES.exe; http://www.abdn.ac.uk/~psy086/dept/; Crawford et al., 2010) and for testing differences in magnitude of condition effects (using DiffBayes_ES.exe; http://www.abdn.ac.uk/~psy086/dept/; Crawford et al., 2010). These single-case analyses take into account the means, standard deviations, sample size, and correlations between tasks in the control sample to determine whether a patient's score (or standardized difference) falls outside the normal range. For space considerations, detailed information about individual patients (precise p values [two-tailed], effect sizes with 95% credible intervals, and correlations between conditions for controls) is provided in the online supplemental materials. The alpha level was .05 for all analyses; however, because this study represents an initial experimental foray into examining feedback and feedforward control in AOS, and sample sizes are unbalanced

---

[4]Error analyses were also conducted. The AOS group made more errors than older controls and the APH group. There were no effects of Condition. See online supplemental materials for further details.

Figure 2. Example of how the average vowel spacing (AVS) measure is derived. The mean M1 × M2 location for each vowel is represented by the symbols. The AVS is the average of the lengths of the broken lines between each possible vowel pair. Thus, a higher AVS indicates greater separation, and therefore greater contrast, between vowels.
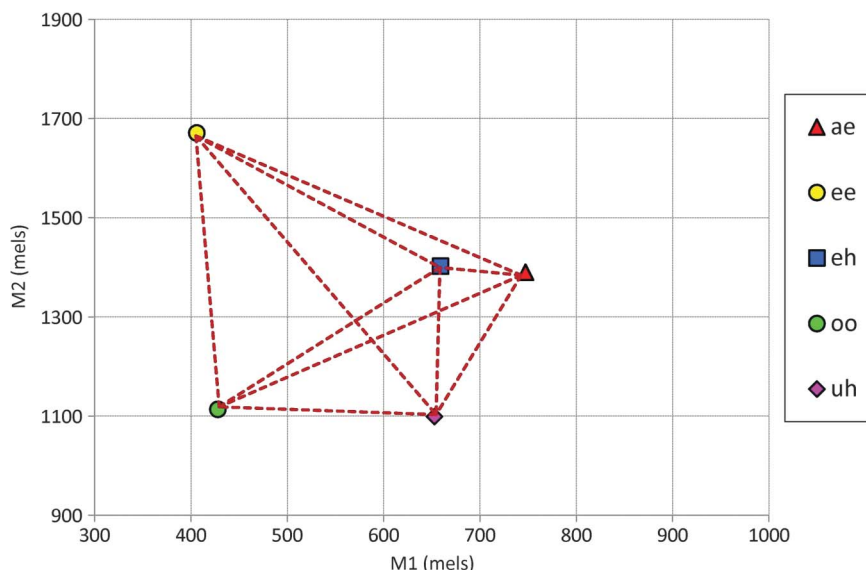
and relatively small, we also report and follow up on trends ($p < .10$). For reported effects, we include generalized eta squared (η^2_G) as an unbiased effect size measure (contrary to partial eta squared, which overestimates effect sizes; Bakeman, 2005; Olejnik & Algina, 2003).

Results

Results from younger adults are presented first, followed by comparisons between older adults and the speaker groups with impairment. Group data are presented in Table 3 (duration, AVS, and dispersion) and Figure 3 (AVS). Data and figures for individual patients (including error rates) are provided in the online supplemental materials.

Younger Adults

Vowel duration. There was a significant effect of Condition, $F(1, 10) = 21.21$, $p = .001$, $\eta^2_G = .49$, with longer duration in the Noise condition (156 ms) than in the Silence condition (135 ms), and a significant effect of Vowel, $F(4, 40) = 44.05$, $p < .001$, $\eta^2_G = .64$, with longer duration for /æ/ than for other vowels and shorter duration for /ʌ/ than for /i/ and /u/. The Condition × Vowel interaction was also significant, $F(4, 40) = 3.94$, $p = .009$, $\eta^2_G = .05$, reflecting that the difference between /ʌ/ and /i/ was significant only in the Silence condition. The Condition effect was present for all vowels.

AVS. There was a main effect of Condition, $F(1, 10) = 29.34$, $p < .001$, $\eta^2_G = .53$, indicating greater AVS in the Silence condition (356 mels) than in the Noise condition (318 mels), and a main effect of Time, $F(2, 20) = 24.67$, $p < .001$, $\eta^2_G = .58$, reflecting greater AVS at 50% (354 mels)

and 80% (350 mels) than at 20% (306 mels). The interaction was also significant, $F(2, 20) = 7.10$, $p = .005$, $\eta^2_G = .04$; the 20% Silence condition differed from the 50% Noise but not from the 80% Noise condition. Note that the Condition effect was present at all time points.

Vowel dispersion. There was an effect of Time, $F(2, 20) = 7.81$, $p = .003$, $\eta^2_G = .03$, with greater dispersion at 50% (40 mels) and 80% (40 mels) than at 20% (35 mels). A significant Vowel × Time interaction, $F(8, 80) = 2.12$, $p = .043$, $\eta^2_G = .04$, reflected greater dispersion at 50% for /ɛ/ than for /i/ and greater dispersion at 80% for /u/ than for /i/.

Older Adults Versus Speakers With Impairment

Vowel duration. A significant Group effect, $F(2, 19) = 7.11$, $p = .005$, $\eta^2_G = .34$, indicated longer vowel duration for the AOS group (182 ms) than for older controls (147 ms) and the APH group (132 ms); the latter two groups did not differ. There was also a Condition effect, $F(1, 19) = 38.75$, $p < .001$, $\eta^2_G = .09$, with longer vowel duration in Noise (165 ms) than in Silence (143 ms) and a main effect of Vowel, $F(4, 76) = 34.18$, $p < .001$, $\eta^2_G = .19$, reflecting longer duration for /æ/ than for all other vowels, which did not differ from each other. The Condition × Vowel interaction was also significant, $F(4, 76) = 3.28$, $p = .016$, $\eta^2_G = .003$, reflecting a difference between /i/ and /ʌ/ (138 vs. 129 ms) in the Silence condition but not in the Noise condition (156 vs. 151 ms). All vowels showed the Condition effect (Noise > Silence).

Individual analyses indicated that four speakers with AOS showed longer vowel duration in the Silence condition (203, 204, 205, 206) and two did so in the Noise condition

Table 3. Group mean (standard deviation) for vowel duration (ms; collapsed across vowels), average vowel spacing (mels), and vowel dispersion (mels; collapsed across vowels) for younger controls (YCON), older controls (OCON), the apraxia of speech (AOS) group, and the aphasia without AOS (APH) group.

Measure	Group	Time point (%)	Condition Silence	Noise	Difference
Duration	YCON (n = 11)		135 (25)	156 (29)	−21
	OCON (n = 12)		136 (17)	157 (23)	−21
	AOS (n = 6)		171 (26)	194 (36)	−23
	APH (n = 4)		120 (23)	143 (24)	−23
AVS	YCON (n = 11)	20	321 (32)	292 (40)	29
		50	373 (31)	334 (44)	39
		80	373 (43)	327 (49)	46
	OCON (n = 12)	20	348 (39)	337 (33)	11
		50	388 (46)	376 (35)	12
		80	356 (44)	341 (35)	14
	AOS (n = 6)	20	343 (84)	306 (74)	37
		50	373 (75)	332 (68)	41
		80	341 (67)	306 (64)	35
	APH (n = 4)	20	364 (83)	341 (83)	23
		50	401 (100)	379 (88)	22
		80	364 (104)	342 (107)	22
Dispersion	YCON (n = 11)	20	35 (9)	36 (8)	−1
		50	41 (11)	40 (9)	1
		80	39 (12)	41 (11)	−2
	OCON (n = 12)	20	36 (8)	35 (7)	−1
		50	39 (10)	37 (10)	−2
		80	41 (10)	40 (10)	−1
	AOS (n = 6)	20	45 (5)	42 (3)	3
		50	51 (10)	43 (5)	8
		80	51 (8)	43 (7)	8
	APH (n = 4)	20	43 (12)	37 (9)	6
		50	45 (14)	32 (7)	13
		80	45 (16)	33 (10)	12

(204, 206; trend for 203). None of the speakers with AOS showed disproportionate Condition effects, except AOS 205, who showed a reverse pattern (longer vowel duration in the Silence condition).

   None of the speakers with aphasia without AOS showed longer duration than age-matched controls (trend for 307 in Silence). APH 304 showed a disproportionate Condition effect compared with the age-matched controls (trend for 307).

   AVS. There was no main effect of Group ($F < 1$), but there was an effect of Condition, $F(1, 19) = 23.79$, $p < .001$, $\eta^2_G = .04$, indicating greater AVS in Silence (363 mels) than in Noise (342 mels), and Time, $F(2, 38) = 9.20$ $p < .001$, $\eta^2_G = .06$, indicating greater AVS at 50% (375 mels) than at 20% (340 mels) and 80% (342 mels). The Group × Condition interaction revealed a trend, $F(2, 19) = 2.79$, $p = .087$, $\eta^2_G = .01$, suggesting that only the AOS group showed a significant Condition effect (Silence: 352 vs. Noise: 315 mels; $p = .004$); this was not evidenced by older controls (364 vs. 351 mels; $p = .365$) or the APH group (376 vs. 354 mels; $p = .318$).

   To address our primary predictions, a separate planned ANOVA involving only the AOS group and age-matched controls indicated significant effects of Condition, $F(1, 16) = 19.25$, $p < .001$, $\eta^2_G = .05$, and Time, $F(2, 32) = 10.08$, $p < .001$, $\eta^2_G = .08$. Critically, the Group × Condition interaction was significant, $F(1, 16) = 4.92$, $p = .042$, $\eta^2_G = .014$, indicating

that the Condition effect was significant in only the AOS group. A similar analysis with only the APH group revealed effects of Condition, $F(1, 14) = 8.72$, $p = .011$, $\eta^2_G = .02$, and Time, $F(2, 28) = 9.23$, $p < .001$, $\eta^2_G = .08$, but no hint of a Group × Condition interaction ($F < 1$).

   Individual analyses indicated that three speakers with AOS differed from controls in both conditions for at least one of the time points; two speakers (200 and 201) showed a lower AVS, and one (205) had a higher AVS. However, of primary interest for our hypotheses was whether the effect of Condition differed in the AOS group compared with control speakers. The pattern indicated by the Group × Condition interaction (larger Condition effect in the AOS group than in older controls) reached significance or trend for at least one time point for three patients (200, 205, and 206) when compared with the older adult control group.

   For the APH speakers, individual analyses indicated that one speaker (304) had higher AVS than controls in both conditions, and one (306) showed lower AVS than controls at the 80% time point in both conditions. None of the speakers showed a disproportionate condition effect at any time point.

   Vowel dispersion. For dispersion, there was a significant main effect of Condition, $F(1, 19) = 12.33$, $p = .002$, $\eta^2_G = .03$, reflecting greater dispersion in the Silence condition (43 mels) than in the Noise condition (38 mels). The

Figure 3. Average vowel spacing (AVS; in mels) across time points for the younger control (YCON) group (top) and for the older control (OCON) group versus the apraxia of speech (AOS) group and aphasia without AOS (APH) group (bottom).
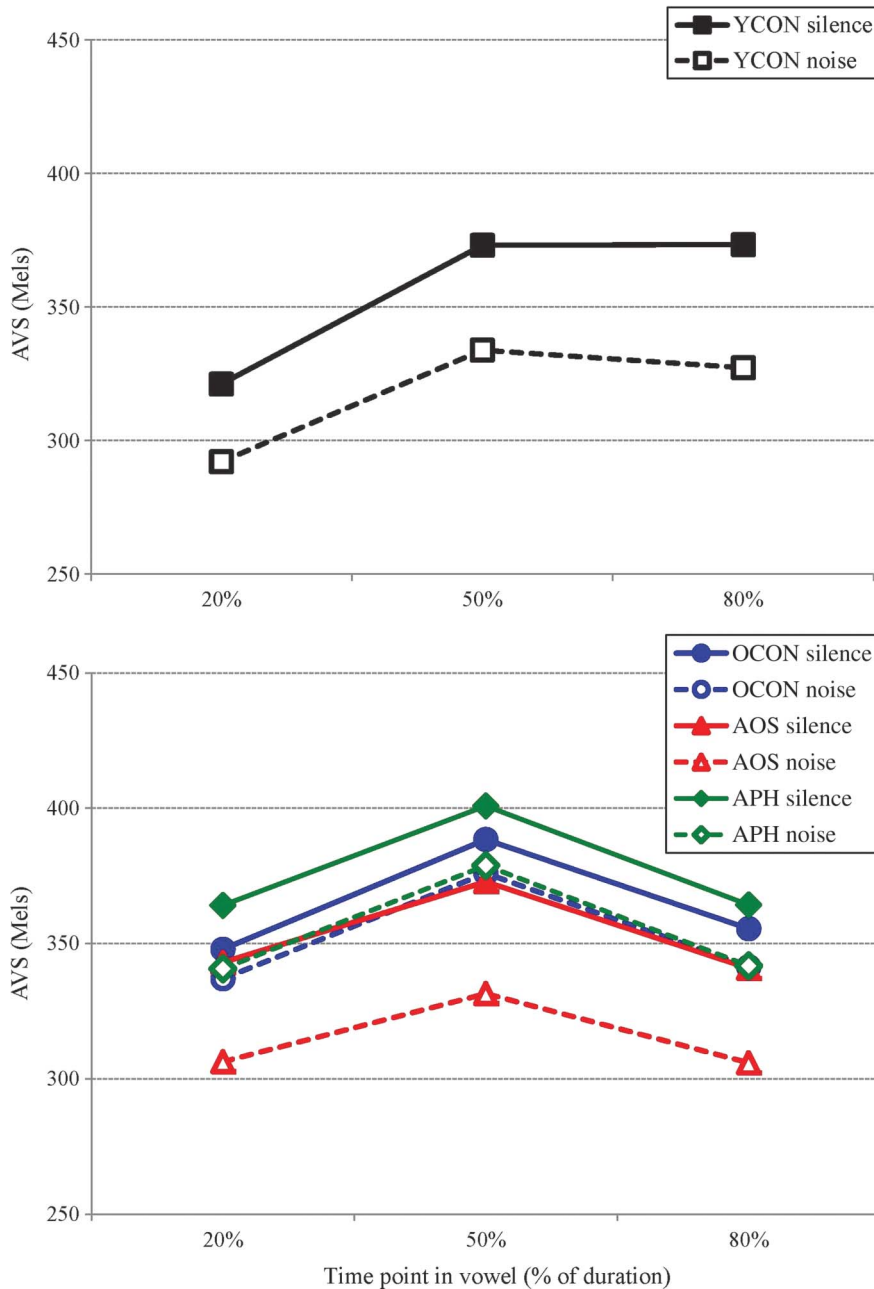
Condition × Time interaction was also significant, $F(2, 38) = 3.87$, $p = .030$, $\eta^2_G = .003$, indicating that dispersion was smaller at 20% (41 mels) than at 50% (45 mels) and 80% (46 mels) in Silence but not in Noise (38, 38, and 39 mels, respectively). A trend toward a Vowel effect, $F(4, 76) = 2.03$, $p = .099$, $\eta^2_G = .03$, suggested greater dispersion for /u/ than for /i/. The Group × Condition interaction reached the level of a trend, $F(2, 19) = 2.62$, $p = .099$, $\eta^2_G = .01$, suggesting that the AOS group had greater dispersion than the control group in Silence (49 vs. 39 mels) but not in Noise (43 vs. 37 mels).

Individual analyses revealed few differences. Of the six AOS speakers, only one (200) demonstrated greater vowel dispersion than controls (at 50% in the Silence condition). Only one speaker demonstrated a disproportionate effect of Condition at 80% (203).

For the APH speakers, individual analyses revealed greater vowel dispersion in the Silence condition for two

speakers (301 at 20%; 304 at 80%). Two of the speakers showed disproportionate Condition effects (at 20% and 50% for 301, and at 80% for 304).

Discussion

This study was designed to experimentally test two hypotheses about AOS in the context of the DIVA model by comparing speech performance with and without self-generated auditory feedback. Before discussing findings from the clinical groups, it is important to briefly review the findings from the younger speaker group. Younger speakers showed no effect of noise masking on vowel dispersion but did show longer vowel durations in the masking condition, consistent with previous studies (Perkell et al., 2007; Rogers et al., 1996). The younger speakers demonstrated a significant reduction in AVS in the masking condition, replicating previous findings (e.g., Lane et al., 2007; Perkell et al., 2007). This demonstrates that our experimental protocol was effective in inducing the masking effect that forms the foundation for the present extension to impaired speech production. These findings suggest that younger speakers monitor and use auditory feedback (in addition to feedforward control) to achieve and maintain segmental contrasts, consistent with the DIVA framework. Although we asked participants to maintain their loudness within habitual levels, we did not measure loudness levels, and thus it is possible that speakers were louder in the Noise condition than in the Silence condition (even if within habitual levels). Such an increase (cf. the Lombard effect; Lane & Tranel, 1971) is typically associated with hyperarticulation (Van Summers et al., 1988) and thus would have resulted in increased AVS. The fact that AVS decreased with masking argues against the interpretation that the condition effect is due to the Lombard effect.

The fact that older control speakers did not show a significant masking effect on vowel contrast was somewhat unexpected. In the context of the DIVA model, a reduced susceptibility to masking of auditory feedback may be attributable to at least two possible factors. First, reduced hearing acuity associated with aging may render the auditory feedback signal less reliable or informative for these speakers, resulting in a greater reliance on feedforward control. This possibility might predict a correlation between hearing thresholds and magnitude of AVS reduction under masking, such that higher thresholds (poorer hearing sensitivity) are associated with a smaller AVS reduction. A post hoc correlation analysis failed to reveal any such correlations at any of the tested frequencies (all p values > .15), although this may reflect in part the small sample size. Future studies with larger sample sizes should take hearing status into account when examining the role of auditory feedback in speech production in older adults and examine how hearing status might affect the Lombard effect. Second, it is possible that feedforward commands are more robust and reliable in older speakers, considering that our older speakers had several decades of additional speaking experience. According to the DIVA model, feedforward commands become more accurate and robust with additional practice, in part because

additional practice provides a wider range of starting positions and corrective commands to be incorporated into the feedforward commands. Although the rate of changes to the feedforward commands is presumably smaller than during speech development, it is possible that differences in feedforward commands can emerge over several decades.

For participants with AOS compared with age-matched controls, our findings revealed the following at the group level: (a) longer vowel duration overall but a comparable vowel-lengthening effect, (b) similar overall acoustic vowel contrast but a disproportionate masking effect, and (c) greater vowel dispersion in the Silence condition.

At a basic level, the longer vowel duration and greater vowel dispersion in the Silence condition are consistent with the diagnosis of AOS in our sample. Longer vowel durations are common in AOS (e.g., Kent & Rosenbek, 1983; Rogers et al., 1996). Greater token-to-token variability in acoustic measures, often interpreted as reflecting speech motor control impairment (e.g., Jacks, 2008), has also been noted in the extant AOS literature (e.g., Seddoh et al., 1996; Whiteside, Grobler, Windsor, & Varley, 2010), though not all studies report greater token-to-token variability (Haley, Ohde, & Wertz, 2001; Jacks et al., 2010).

With respect to our main question regarding the underlying nature of the deficit in AOS, the data revealed a disproportionate masking effect for the AOS group compared with age-matched controls. Although this interaction was small and only reached the level of a trend in the omnibus ANOVA with the three unbalanced groups, the planned comparison clearly revealed the pattern of greater AVS reduction in AOS speakers than in age-matched controls at the group level. A disproportionate effect of noise masking on acoustic vowel contrast is consistent with and predicted by the FF hypothesis but inconsistent with the FB hypothesis, which predicts an increase or no change in AVS with noise masking in AOS.

According to the FF hypothesis, speakers with AOS have impaired feedforward commands (Jacks, 2008), which results in a greater reliance on auditory feedback control to achieve and maintain segmental contrasts. When auditory feedback is not available, segmental contrast (measured here in terms of AVS) was predicted to show a larger decrease, relative to age-matched control speakers. This is precisely the pattern that was observed at the group level. However, although these group-level data provide some support for the FF hypothesis, not all speakers demonstrated the group pattern. Although individual data are not always reported or discussed (e.g., Aichert & Ziegler, 2004; Deger & Ziegler, 2002; Seddoh et al., 1996), observations that not all participants with AOS display the group pattern are not uncommon (e.g., Aichert & Ziegler, 2012; Haley et al., 2001; Mailend & Maas, 2013). The individual analyses in the present study indicated a disproportionate reduction in AVS with masking for at least two speakers with AOS (200 and 205), with a trend for a third (206). Based on the logic of the theoretical framework, these findings can be viewed as positive evidence of impaired feedforward control in these individuals and an ability to use auditory feedback

to enhance vowel contrasts. For the other speakers with AOS, the group pattern was present numerically but not statistically, and as such, we cannot conclude that their feedforward systems are impaired (though it also does not mean that their feedforward systems are intact). However, we can exclude a feedback deficit for at least two of these individuals (201 and 204) because they did not show an unchanged AVS with decreased dispersion; for AOS 203, the lack of a disproportionate reduction in AVS occurred in the context of a disproportionate reduction in vowel dispersion, suggestive of the pattern predicted by the FB hypothesis. Examination of the clinical and demographic profiles of our participants with AOS did not reveal any clear factor that might account for the interindividual variability in pattern observed in this study.

The fact that not all speakers with AOS in this study demonstrated the group pattern to a significant extent in the individual analyses may be due to the relatively small control group (i.e., limited power) or to the possibility that different profiles of impairment exist among the group of speakers broadly diagnosed as having AOS. In fact, models such as the DIVA model, in presenting a more sophisticated, multicomponent view of speech motor planning, are highly compatible with the existence of different impairment profiles ("subtypes").

Given the high variation of lesion sites and extents in the AOS group, combined with the complexity of the speech production system, an expectation of a single underlying core impairment shared by all individuals with AOS may be too simplistic. By analogy, the diagnosis of aphasia also covers a wide range of underlying language impairments. Although opinions differ on how to parse the interindividual differences among people with aphasia (e.g., Ardila, 2010; Caplan, 1993; Howard, Swinburn, & Porter, 2010; Schwartz, 1984), the notion that generalizations about aphasia hold for all people with aphasia (or even for a particular aphasia type; cf. Berndt, Mitchum, & Haendiges, 1996; Schwartz, 1984) is also tenuous. Some rather generic statements may hold (such as all individuals with aphasia have some degree of word-finding difficulty), although even in that case, the reasons may differ among patients (e.g., Kohn & Goodglass, 1985). Thus, in aphasiology it is uncontroversial to assume that not all people with aphasia share the same single underlying deficit; different profiles of language impairment are expected given the complexity of the language system. A more detailed understanding of the language system and the factors that influence language processing has also led to development of more targeted approaches to the study and assessment of aphasia (e.g., psycholinguistic approaches; Caplan, 1993; Howard et al., 2010). Given that diagnosis of AOS is currently based on a clinical expert's (generally unquantified) judgment of presence and absence of certain speech features observed in a variety of speech tasks, current diagnostic standards may lack the precision to identify potential subtypes or profiles of AOS.

Although speculation about AOS subtypes is not new (e.g., Kent, 1991; Luria & Hutton, 1977; Mlcoch & Noll,

1980), the development of sophisticated models of speech motor planning offers principled frameworks for hypothesizing different loci of speech motor planning deficits and possibly experimental approaches to test these hypotheses. The present study represents an initial foray into developing and testing such model-driven hypotheses about underlying speech motor planning impairments. Our findings indicate that for some speakers with AOS, the feedforward system is impaired, whereas for other speakers diagnosed with AOS, the problem may lie in different components of the system (e.g., the retrieval of motor programs from the SSM; cf. Mailend & Maas, 2013). Thus, it may be more appropriate to consider the FF hypothesis as a hypothesis of a possible subtype of AOS that applies to some, but not all, individuals with AOS.

The notion that speakers with AOS rely to a greater extent on auditory feedback during normal speech has been addressed previously by Rogers et al. (1996), who examined the effects of white noise masking on vowel duration. They argued that speakers with AOS may produce longer vowels due to a reliance on auditory feedback control. They reasoned that removing auditory feedback should invalidate this strategy and therefore result in shorter vowels under masking conditions. Instead, Rogers et al. observed vowel lengthening with masking noise, both in their three speakers with AOS and in the three control speakers, without any indication of exaggerated or reduced lengthening effects in the AOS speakers. The present study replicates this finding, with a larger sample size and similar speech materials: Both groups produced longer vowels under masking conditions, with no indication of a disproportionate effect for AOS speakers. Rogers et al. suggested that longer vowel duration in speakers with AOS is not due to monitoring auditory information and instead may reflect additional planning time for retrieving, programming, or initiating subsequent segments. Although this may indeed be the case (cf. Mailend & Maas, 2013), it is also possible that a greater reliance on auditory feedback control is not reflected in duration or that vowels are prolonged in an effort to obtain somatosensory feedback to control speech articulation (Fucci, Crary, Warren, & Bond, 1977). Instead, spectral measures of segmental contrast, as in the present study, provide a more direct way to examine the role of auditory feedback in maintaining segmental contrast in speech production in AOS.

The interpretation of greater reliance on feedback control in some speakers with AOS is consistent with and complements the findings of Jacks (2008), who used a different research strategy. In particular, Jacks compared vowel contrast with and without a bite block in speakers with AOS and control speakers. Bite blocks essentially force greater reliance on feedback control because they introduce errors into the speech generated by the feedforward commands. Speakers with AOS demonstrated smaller vowel contrast than controls (possibly due to impaired feedforward commands), but both groups showed comparable compensations for the bite blocks, suggesting that speakers with AOS have intact feedback control. Thus, whereas Jacks's paradigm essentially forced greater reliance on feedback by perturbing speech

with bite blocks and suggested that speakers with AOS are able to use feedback control, the present paradigm forced greater reliance on feedforward control by masking auditory feedback and suggested that at least some speakers with AOS exhibit a deficit in the feedforward system.

The present study also extends previous work by including speakers with aphasia without AOS because there are indications in the literature that even those speakers may have subtle speech motor control deficits (e.g., McNeil, Liss, Tseng, & Kent, 1990). Although the sample was small, and replication is needed, there was no indication for a disproportionate vowel contrast reduction in either the group analysis or for any individual speaker, suggesting that the effect observed in some of our speakers with AOS is not attributable to aphasia. Two of the APH speakers did show greater vowel dispersion in the Silence condition (cf. Ryalls, 1986), which may indicate a greater contribution of auditory feedback in vowel articulation for these speakers. However, neither of these individuals demonstrated a reduced contrast when auditory feedback was removed, suggesting that a greater contribution of auditory feedback may not reflect a compensation for impaired feedforward control. The exact nature of this greater dispersion in the presence of self-generated auditory feedback is unclear. One possibility is that the lesions in these patients interfere with the processing of auditory feedback, which in turn might interfere with the ongoing motor commands, thereby increasing variability (Ryalls, 1986).

Although the DIVA model does not make specific predictions about individuals with aphasia without AOS, a recently proposed hierarchical state feedback control model (Hickok, 2012, 2014; Hickok, Houde, & Rong, 2011) suggests that conduction aphasia reflects an impairment of the linkage between auditory targets and motor codes. This may disrupt an internal error monitoring process, resulting in sound-based speech errors and an inability to correct such errors when they are perceived (through a comparison of overt acoustic feedback with correctly activated auditory targets; Hickok et al., 2011). Given this conceptualization of conduction aphasia as involving essentially an auditory feedback control deficit (Hickok et al., 2011, p. 416; cf. our FB hypothesis), one might predict that individuals with conduction aphasia would show unchanged vowel contrast with masking and a disproportionate reduction of token-to-token variability. Neither of the two speakers with conduction aphasia in our sample (304 and 307) showed this pattern. Participant 304 did show a disproportionate reduction in vowel dispersion, but this was accompanied by a decrease in AVS rather than an increase. These findings must of course be viewed with caution, given the small sample size and the fact that the study was not designed to test hypotheses about aphasia. It is possible that none of our patients had damage to the Sylvian parietal temporal area at the temporoparietal junction (area Spt), the area presumed to underlie integration of auditory and motor codes in the hierarchical state feedback control model (Hickok et al., 2011). The available lesion data are not sufficiently detailed to address this possibility.

Taken together, given the relatively small sample size and the notorious variability in AOS (e.g., Bartle-Meyer & Murdoch, 2010; Seddoh et al., 1996; Whiteside et al., 2010), our findings provide some support for a deficit in feedforward control in at least some speakers with AOS, with the caveats noted herein. Clearly, future research is needed to replicate these findings and further explore the factors that affect response to masking. For example, although our protocol included a hearing screening, we did not examine more specific auditory perceptual abilities (Perkell et al., 2004) or somatosensory acuity (Ghosh et al., 2010) to determine how individuals' perceptual abilities might modulate or compensate for the effects of underlying impairments. In addition, because of some excluded trials (e.g., due to errors, measurement difficulties), we were unable to compute AVS for each block of trials separately and instead computed AVS on all available trials across blocks. This made it impossible to address any potential carryover effects between blocks. However, we did conduct an additional analysis in which we included Half as a factor (first eight blocks vs. second eight blocks). There were no effects or interactions with Half for any of the dependent measures, indicating that the effects of interest did not differ from the first half to the second half of the experiment. Future studies could address this issue further. Perhaps most important, given that not all participants with AOS demonstrated evidence for a feedforward deficit (nor for a feedback deficit), future research should investigate additional model-driven hypotheses about the nature of speech motor planning impairments in individuals diagnosed clinically with AOS. Inclusion of detailed lesion information will enable exploration of brain–behavior relationships.

Ultimately, sophisticated and detailed frameworks of speech motor control such as the DIVA model may lay the groundwork for a process-oriented approach to assessment that goes beyond determining whether an individual has a "speech motor planning" impairment (i.e., AOS) and allows researchers and clinicians to identify more specifically *which parts* of the speech motor planning process are impaired, akin to the psycholinguistic approach to assessment of aphasia (e.g., Caplan, 1993). Computational modeling simulations of hypothesized impairments will also be important in order to provide a deeper understanding of the possible interactions among components of the system and generate quantitative predictions for individual speakers. We are currently exploring such simulations with respect to AOS, similar to computational simulations for childhood apraxia of speech (Terband, Maassen, Guenther, & Brumberg, 2009, 2014) and stuttering (Civier, Tasko, & Guenther, 2010). Such advancements in our understanding and assessment of impairments underlying AOS may also facilitate development of validated and operationalized diagnostic markers (and thereby improve agreement on diagnosis) as well as more targeted intervention approaches.

Conclusions

In this study, we experimentally manipulated the availability of auditory feedback during speech production

in speakers with and without AOS to evaluate two hypotheses concerning the nature of AOS derived from the DIVA model. The disproportionate reduction in vowel contrast with noise masking lends support to the hypothesis that for some speakers with AOS, the impairment involves a deficit in feedforward control, consistent with recent literature (e.g., Jacks, 2008), although this interpretation must be taken with caution given the small sample size, small effect sizes, and interindividual variability that might suggest the existence of AOS subtypes. No disproportionate reduction of vowel contrast was evident for speakers with aphasia without AOS. Clearly, further research is needed to replicate these findings and further explore the possibility for subtypes of AOS reflecting different underlying speech motor planning impairments. Finally, our findings add to the growing body of evidence for the utility of the DIVA model in generating and testing hypotheses about speech disorders (e.g., stuttering: Civier et al., 2010; childhood apraxia of speech: Terband et al., 2009; and AOS: Jacks, 2008; Mailend & Maas, 2013).

Acknowledgments

References

Aichert, I., & Ziegler, W. (2004). Syllable frequency and syllable structure in apraxia of speech. *Brain and Language, 88,* 148–159.

Aichert, I., & Ziegler, W. (2012). Word position effects in apraxia of speech: Group data and individual variation. *Journal of Medical Speech-Language Pathology, 20*(4), 7–11.

American Speech-Language-Hearing Association. (1997). *Guidelines for audiologic screening* [Guidelines]. Available from http://www.asha.org/policy

Ardila, A. (2010). A proposed reinterpretation and reclassification of aphasic syndromes. *Aphasiology, 24,* 363–394.

Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior Research Methods, 37,* 379–384.

Ballard, K. J., & Robin, D. A. (2007). Influence of continual biofeedback on jaw pursuit-tracking in healthy adults and in adults with apraxia plus aphasia. *Journal of Motor Behavior, 39,* 19–28.

Bartle-Meyer, C. J., & Murdoch, B. E. (2010). A kinematic investigation of anticipatory lingual movement in acquired apraxia of speech. *Aphasiology, 24,* 623–642.

Berndt, R. S., Mitchum, C. C., & Haendiges, A. N. (1996). Comprehension of reversible sentences in "agrammatism": a meta-analysis. *Cognition, 58,* 289–308.

Boersma, P., & Weenink, D. (2008). Praat: Doing Phonetics by Computer (Version 5.0.17) [Computer program]. Retrieved from http://www.praat.org/

Bohland, J. W., Bullock, D., & Guenther, F. H. (2010). Neural representations and mechanisms for the performance of simple speech sequences. *Journal of Cognitive Neuroscience, 22,* 1504–1529.

Cai, S., Ghosh, S. S., Guenther, F. H., & Perkell, J. S. (2011). Focal manipulations of formant trajectories reveal a role of auditory feedback in the online control of both within-syllable and between-syllable speech timing. *Journal of Neuroscience, 31,* 16483–16490.

Caplan, D. (1993). Toward a psycholinguistic approach to acquired neurogenic language disorders. *American Journal of Speech-Language Pathology, 2*(1), 59–83.

Civier, O., Tasko, S. M., & Guenther, F. H. (2010). Overreliance on auditory feedback may lead to sound/syllable repetitions: Simulations of stuttering and fluency-inducing conditions with a neural model of speech production. *Journal of Fluency Disorders, 35,* 246–279.

Crawford, J. R., Garthwaite, P. H., & Porter, S. (2010). Point and interval estimates of effect sizes for the case-controls design in neuropsychology: Rationale, methods, implementations, and proposed reporting standards. *Cognitive Neuropsychology, 27,* 245–260.

Dabul, B. (2000). *The Apraxia Battery for Adults–Second Edition.* Austin, TX: Pro-Ed.

Deal, J., & Darley, F. L. (1972). The influence of linguistic and situational variables on phonemic accuracy in apraxia of speech. *Journal of Speech and Hearing Research, 15,* 639–653.

Deger, K., & Ziegler, W. (2002). Speech motor programming in apraxia of speech. *Journal of Phonetics, 30,* 321–335.

Duffy, J. R. (2005). *Motor speech disorders: Substrates, differential diagnosis, and management* (2nd ed.). St. Louis, MO: Mosby-Yearbook.

Fucci, D., Crary, M., Warren, J., & Bond, Z. (1977). Interaction between auditory and oral sensory feedback in speech regulation. *Perceptual & Motor Skills, 45,* 123–129.

Ghosh, S. S., Matthies, M. L., Maas, E., Hanson, A., Tiede, M., Ménard, L., & Perkell, J. S. (2010). An investigation of the relation between sibilant production and somatosensory and auditory acuity. *The Journal of the Acoustical Society of America, 128,* 3079–3087.

Goodglass, H., Kaplan, E., & Barresi, B. (2000). *Boston Diagnostic Aphasia Examination–Third Edition.* Philadelphia, PA: Lippincott Williams & Wilkins.

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,* 280–301.

Guenther, F. H., Hampson, M., & Johnson, D. (1998). A theoretical investigation of reference frames for the planning of speech movements. *Psychological Review, 105,* 611–633.

Haley, K. L., Jacks, A., de Riesthal, M., Abou-Khalil, R., & Roth, H. L. (2012). Toward a quantitative basis for assessment and diagnosis of apraxia of speech. *Journal of Speech, Language, and Hearing Research, 55,* S1502–S1517.

Haley, K. L., Ohde, R. N., & Wertz, R. T. (2001). Vowel quality in aphasia and apraxia of speech: Phonetic transcription and formant analyses. *Aphasiology, 15,* 1107–1123.

Hickok, G. (2012). Computational neuroanatomy of speech production. *Nature Reviews Neuroscience, 13,* 135–145.

Hickok, G. (2014). The architecture of speech production and the role of the phoneme in speech processing. *Language, Cognition, and Neuroscience, 29,* 2–20.

Hickok, G., Houde, J., & Rong, F. (2011). Sensorimotor integration in speech processing: Computational basis and neural organization. *Neuron, 69,* 407–422.

Hillis, A. E., Work, M., Barker, P. B., Jacobs, M. A., Breese, E. L., & Maurer, K. (2004). Re-examining the brain regions crucial for orchestrating speech articulation. *Brain, 127,* 1479–1487.

Honda, M., Fujino, A., & Kaburagi, T. (2002). Compensatory responses of articulators to unexpected perturbation of the palate shape. *Journal of Phonetics, 30,* 281–302.

Houde, J. F., & Jordan, M. I. (1998, February 20). Sensorimotor adaptation in speech production. *Science, 279,* 1213–1216.

Howard, D., Swinburn, K., & Porter, G. (2010). Putting the CAT out: What the Comprehensive Aphasia Test has to offer. *Aphasiology, 24,* 56–74.

Jacks, A. (2008). Bite block vowel production in apraxia of speech. *Journal of Speech, Language, and Hearing Research, 51,* 898–913.

Jacks, A., Mathes, K. A., & Marquardt, T. P. (2010). Vowel acoustics in adults with apraxia of speech. *Journal of Speech, Language, and Hearing Research, 53,* 61–74.

Kelso, J. A. S., & Tuller, B. (1983). "Compensatory articulation" under conditions of reduced afferent information: A dynamic formulation. *Journal of Speech and Hearing Research, 26,* 217–224.

Kent, R. D. (1991). Research needs in the assessment of speech motor disorders. In J. A. Cooper (Ed.), *Assessment of speech and voice production: Research and clinical applications* (pp. 17–29). Bethesda, MD: National Institute on Deafness and Other Communication Disorders.

Kent, R. D., Kent, J. F., Weismer, G., & Duffy, J. R. (2000). What dysarthrias can tell us about the neural control of speech. *Journal of Phonetics, 28,* 273–302.

Kent, R. D., & Rosenbek, J. C. (1983). Acoustic patterns of apraxia of speech. *Journal of Speech and Hearing Research, 26,* 231–249.

Kertesz, A. (2006). *The Western Aphasia Battery–Revised.* San Antonio, TX: Pearson.

Kjellberg, A. (1990). Subjective, behavioral, and psychophysiological effects of noise. *Scandinavian Journal of Work and Environmental Health, 16,* 29–38.

Kohn, S. E., & Goodglass, H. (1985). Picture-naming in aphasia. *Brain and Language, 24,* 266–283.

Lane, H., Denny, M., Guenther, F. H., Matthies, M. L., Ménard, L., Perkell, J. S., ... Zandipour, M. (2005). Effects of bite blocks and hearing status on vowel production. *The Journal of the Acoustical Society of America, 118,* 1636–1646.

Lane, H., Matthies, M. L., Guenther, F. H., Denny, M., Perkell, J. S., Stockmann, E., ... Zandipour, M. (2007). Effects of short- and long-term changes in auditory feedback on vowel and sibilant contrasts. *Journal of Speech, Language, and Hearing Research, 50,* 913–927.

Lane, H., & Tranel, B. (1971). The Lombard sign and the role of hearing in speech. *Journal of Speech and Hearing Research, 14,* 677–709.

Leung, M.-T., & Ciocca, V. (2011). The effects of tongue loading and auditory feedback on vowel production. *The Journal of the Acoustical Society of America, 129,* 316–325.

Luria, A. R., & Hutton, J. T. (1977). A modern assessment of the basic forms of aphasia. *Brain and Language, 4,* 129–151.

Maas, E., Robin, D. A., Wright, D. L., & Ballard, K. J. (2008). Motor programming in apraxia of speech. *Brain and Language, 106,* 107–118.

Mailend, M.-L., & Maas, E. (2013). Speech motor programming in apraxia of speech: Evidence from a delayed picture-word interference task. *American Journal of Speech-Language Pathology, 22,* S380–S396.

McNeil, M. R., Liss, J. M., Tseng, C.-H., & Kent, R. D. (1990). Effects of speech rate on the absolute and relative timing of apraxic and conduction aphasic sentence production. *Brain and Language, 38,* 135–158.

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* (2nd ed., pp. 249–268). New York, NY: Thieme.

Mlcoch, A. G., & Noll, J. D. (1980). Speech production models as related to the concept of apraxia of speech. In N. J. Lass (Ed.), *Speech and language: Advances in basic research and practice* (Vol. 4, pp. 201–239). New York, NY: Academic Press.

Morrell, C. H., Gordon-Salant, S., Pearson, J. D., Brant, L. J., & Fozard, J. L. (1996). Age- and gender-specific reference ranges for hearing level and longitudinal changes in hearing level. *The Journal of the Acoustical Society of America, 100,* 1949–1967.

Nilsson, M., Soli, S. D., & Sullivan, J. A. (1994). Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise. *The Journal of the Acoustical Society of America, 95,* 1085–1099.

Olejnik, S., & Algina, J. (2003). Generalized eta and omega squared statistics: Measures of effect size for some common research designs. *Psychological Methods, 8,* 434–447.

Perkell, J. S. (2012). Movement goals and feedback and feedforward control mechanisms in speech production. *Journal of Neurolinguistics, 25,* 382–407.

Perkell, J. S., Denny, M., Lane, H., Guenther, F. H., Matthies, M. L., Tiede, M., ... Burton, E. (2007). Effects of masking noise on vowel and sibilant contrasts in normal-hearing speakers and postlingually deafened cochlear implant users. *The Journal of the Acoustical Society of America, 121,* 505–518.

Perkell, J. S., Guenther, F. H., Lane, H., Matthies, M. L., Stockmann, E., Tiede, M., & Zandipour, M. (2004). The distinctness of speakers' productions of vowel contrasts is related to their discrimination of the contrasts. *The Journal of the Acoustical Society of America, 116,* 2338–2344.

Purcell, D. W., & Munhall, K. G. (2006). Compensation following real-time manipulation of formants in isolated vowels. *The Journal of the Acoustical Society of America, 119,* 2288–2297.

Reinfeldt, S., Östli, P., Håkansson, B., & Stenfelt, S. (2010). Hearing one's own voice during phoneme vocalization—Transmission by air and bone conduction. *The Journal of the Acoustical Society of America, 128,* 751–762.

Rogers, M. A., Eyraud, R., Strand, E. A., & Storkel, H. (1996). The effects of noise masking on vowel duration in three patients with apraxia of speech and a concomitant aphasia. *Clinical Aphasiology, 24,* 83–96.

Ryalls, J. H. (1986). An acoustic study of vowel production in aphasia. *Brain and Language, 29,* 48–67.

Schwartz, M. F. (1984). What the classical aphasia syndromes can't do for us, and why. *Brain and Language, 21,* 3–8.

Seddoh, S. A. K., Robin, D. A., Sim, H. S., Hageman, C., Moon, J. B., & Folkins, J. W. (1996). Speech timing in apraxia of speech versus conduction aphasia. *Journal of Speech and Hearing Research, 39,* 590–603.

Shiller, D. M., Sato, M., Gracco, V. L., & Baum, S. R. (2009). Perceptual recalibration of speech sounds following speech motor learning. *The Journal of the Acoustical Society of America, 125,* 1103–1113.

Square, P. A., Darley, F. L., & Sommers, R. K. (1981). Speech perception among patients demonstrating apraxia of speech, aphasia, and both disorders. *Clinical Aphasiology, 9,* 83–88.

Terband, H., Maassen, B., Guenther, F. H., & Brumberg, J. (2009). Computational neural modeling of speech motor control in childhood apraxia of speech (CAS). *Journal of Speech, Language, and Hearing Research, 52,* 1595–1609.

Terband, H., Maassen, B., Guenther, F. H., & Brumberg, J. (2014). Auditory-motor interactions in pediatric motor speech disorders: Neurocomputational modeling of disordered development. *Journal of Communication Disorders, 47,* 17–33.

Tourville, J. A., Reilly, K. J., & Guenther, F. H. (2008). Neural mechanisms underlying auditory feedback control of speech. *NeuroImage, 39,* 1429–1443.

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., pp. 3–18). New York, NY: Thieme.

Van Summers, W., Pisoni, D. B., Bernacki, R. H., Pedlow, R. I., & Stokes, M. A. (1988). Effects of noise on speech production: Acoustic and perceptual analyses. *The Journal of the Acoustical Society of America, 84,* 917–928.

Villacorta, V., Perkell, J. S., & Guenther, F. H. (2007). Sensorimotor adaptation to feedback perturbations of vowel acoustics and its relation to perception. *The Journal of the Acoustical Society of America, 122,* 2306–2319.

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.

Whiteside, S. P., Grobler, S., Windsor, F., & Varley, R. (2010). An acoustic study of vowels and coarticulation as a function of utterance type: A case of acquired apraxia of speech. *Journal of Neurolinguistics, 23,* 145–161.

Ziegler, W. (2002). Psycholinguistic and motor theories of apraxia of speech. *Seminars in Speech and Language, 23,* 231–244.