Effects of Feedback Frequency and Timing on Acquisition, Retention, and Transfer of Speech Skills in Acquired Apraxia of Speech

Shannon N. Auerstrom Hula
Donald A. Robin
Edwin Maas
San Diego State University/University of California San Diego

Kirrie J. Ballard
University of Iowa, Iowa City

Richard A. Schmidt
University of California Los Angeles

Purpose: Two studies examined speech skill learning in persons with apraxia of speech (AOS). Motor-learning research shows that delaying or reducing the frequency of feedback promotes retention and transfer of skills. By contrast, immediate or frequent feedback promotes temporary performance enhancement but interferes with retention and transfer. These principles were tested in the context of a common treatment for AOS.

Method: Two studies (N = 4, N = 2) employed single-subject treatment designs to examine acquisition and retention of speech skills in adults with AOS under different feedback conditions.

Results: Reduced-frequency or delayed feedback enhanced learning in 3 participants with AOS. Feedback manipulation was not an influential variable in 3 other cases in which stimulus-complexity effects may have masked treatment effects.

Conclusions: These findings demonstrate that individuals with AOS can benefit from structured intervention. They provide qualified support for reduction and delay of feedback, although interaction with other factors such as stimulus complexity or task difficulty needs further exploration. This study adds to the growing body of literature investigating the use of principles of motor learning in treating AOS and provides impetus for consideration of pre-treatment variables that affect outcome in treatment studies.

KEY WORDS: apraxia of speech, apraxia treatment, feedback, principles of motor learning

The purpose of the current experiments was to explore the effects of feedback frequency and timing on the acquisition, retention, and transfer of speech skills in persons with apraxia of speech (AOS). These data extend work on principles of motor learning in limb motor learning to the treatment of speech as well as provide more data on treatment efficacy in AOS. The theoretical motivation for this work has its genesis in the schema theory of motor control and learning (Schmidt, 1975). However, other views of motor learning and programming (e.g., Li & Wright, 2000; Shea & Wulf, 2005) that can be integrated with schema theory are also compatible with our theoretical approach. AOS is best considered a disorder of motor control and, in particular, one of motor programming (McNeil, Robin, & Schmidt, 1997). We use the term motor programming to refer to the set of processes that specify “a representation that, when initiated, results in the production of a coordinated movement sequence” (Schmidt & Lee, 2005, p. 466). Because principles
of motor learning are thought to operate, in part, at the programming level, this group of patients is ideal to test the application of these principles to the treatment of speech. The theoretical framework will first be outlined, followed by discussion of its recent application to motor speech learning and AOS.

Theoretical Framework

Motor learning refers to a “set of internal processes associated with practice or experience leading to relatively permanent changes in the capability for movement” (Schmidt & Lee, 2005, p. 302). The schema theory of motor control and learning proposed by Schmidt (1975) assumes that the brain stores “generalized” motor programs (GMPs) that represent the relative timing and relative force of muscle commands necessary for carrying out members of a class of movements. Given particular response specifications, the processing mechanism selects the parameters, or details, of the movement to be executed (Shea & Wulf, 2005). The parameters assigned to a given GMP specify the absolute timing and absolute force of muscle contractions in the chosen effectors. It is assumed that the specification of GMPs and the selection of parameters are shaped and refined during motor learning.

Schema Theory

Although other theories of motor control exist (e.g., Kelso & Tuller, 1981; Saltzman & Munhall, 1989; Thelen & Smith, 1994), schema theory provides the theoretical framework for this program of research because of its emphasis on motor learning and its influence on the development of specific principles of motor learning such as those addressed in this paper. Other authors (Shea & Wulf, 2005) have recently advocated a reconceptualization of the GMP as a “scalable response structure” (SRS) and emphasize processing mechanisms instead of schemata, but the terms “GMP” and “schemata” have been used here to maintain terminological consistency with the literature from which this article draws.

According to schema theory, four types of information are stored after a movement is executed. This information includes the initial conditions (task and environment conditions prior to movement production), the parameters that are assigned to the GMP, the outcome of the movement in terms of the environmental goal, and the sensory consequences of the movement. In order to learn new skills, reorganize older skills to be performed at more challenging levels, and presumably to re-learn skills that have been lost, the performer must develop abstract relationships between these pieces of information upon completion of the movement. The first of these relationships, the recall schema, involves the relationship among past parameters, the initial conditions, and the movement outcomes produced by their combinations. In future trials, when the initial conditions and desired outcomes are noted, the recall schema serves to select the parameters most appropriate for achieving the movement goal and applies them to the GMP. The second abstract relationship is the recognition schema. This is the relationship among the past initial conditions, the past environmental outcomes, and the past sensory consequences of those movements. When a performer notes the initial conditions and desired outcome before a movement, he or she can then estimate (anticipate) the sensory consequences of that movement. These expected sensory consequences are then compared to the actual feedback produced in order to evaluate the movement after its execution to detect error. Under the framework of schema theory, learning consists of refining/strengthening these two schemata through experience.

Schema theory makes several predictions about how development of GMPs and schemata might be affected by specific conditions present during practice (see Schmidt & Lee, 2005; Shea & Wulf, 2005). These conditions define how new skills are practiced and the type and frequency of external, or augmented, feedback that is provided to the learner. Through numerous studies of motor learning in the limbs, a set of motor learning principles has been identified that differentiates between variables that enhance performance temporarily and those that bring about robust long-term learning (for reviews, see Maas et al., 2008; Schmidt & Lee, 2005). This distinction between performance and long-term learning then requires that one measures short-term changes in performance, from training session to training session, as well as long-term retention of trained skills after termination of training and transfer of trained skills to related but untrained skills.

Augmented Feedback

Although there are a number of principles of motor learning, we focus here on those involving provision of augmented feedback, a ubiquitous component of motor speech treatments. Due to the difficulty of studying the effects of intrinsic sources of outcome information in humans, researchers have developed paradigms for manipulating extrinsic, “augmented” feedback as a means of deducing the operations of intrinsic feedback in naturalistic contexts. Studies in limb motor learning have shown that increased frequency of external knowledge of results (KR) feedback promotes parameter learning (e.g., Wulf, Schmidt, & Deubel, 1993). However, the
provision of too much external KR feedback appears to degrade GMP learning, suggesting that reduced availability of external outcome information is important for promoting performers’ learning of the core features of a movement pattern, or GMP (Wulf, Lee, & Schmidt, 1994). To date, studies have found that reduced KR is either equally or more effective in promoting learning (e.g., Lee, White, & Carnahan, 1990; Sparrow & Summers, 1992; Weinstein & Schmidt, 1990). No studies have reported a superior effect of 100% KR over reduced KR on learning.

In addition to frequency of KR feedback, the temporal locus of feedback is an important determinant of the availability of external outcome information. Researchers have examined three intervals: feedback delay interval (time between participant’s production and provision of feedback), postfeedback delay interval (time elapsed between feedback and next stimulus presentation), and intertrial interval (time between two successive trials; Schmidt & Lee, 2005). Animal studies have suggested that delaying a reward after response to a stimulus has a detrimental effect on conditioning (e.g., Perin, 1943; Skinner, 1936). Based on this, researchers became interested in evaluating the effects of feedback delay on human motor learning (see Salmoni, Schmidt, & Walter, 1984, for review). Although most studies have revealed null or inconsistent effects of feedback delay on human motor skills learning (e.g., Becker, Mussina, & Persons, 1963; Koch & Dorfman, 1979; Mulder & Hulstijn, 1985; Weltens & de Bot, 1984), most have focused on temporary changes in performance rather than long-term retention and transfer of motor skills. Swinnen, Schmidt, Nicholson, and Shapiro (1990) examined the effect of a short feedback delay on acquisition and retention of a bat-swing motor skill. They concluded that instantaneous feedback initially supported acquisition of the behavior but at a certain point began to impede the continued improvement during acquisition. Furthermore, it interfered with retention of the trained skill at 10 min after training and, more dramatically, 2 days after training. This trend persisted on a 4-month retention test, although the group difference was no longer significant. More recently, Anderson, Magill, Sekiya, and Ryan (2005) reported that delayed feedback (i.e., feedback given after two intervening trials) resulted in less accurate acquisition performance of an unfamiliar aiming behavior but stronger retention after a 24-hr delay. The size of this difference was moderate, although it did not reach significance. Additionally, the decline in performance from acquisition to retention (at 1 min and at 24 hr postacquisition) was smaller for the delayed feedback group than for the immediate feedback group, and the delayed group reported using a greater number and variety of intrinsic feedback sources during practice.

The vast majority of limb motor learning studies have tested healthy individuals. Several studies have extended the work on principles of motor learning to patients with neurological disease in re-learning limb control (Goodgold-Edwards & Cermak, 1990; Hanlon, 1996; Jarus, 1994; Sabari, 1991; Stevans & Hall, 1998). In general, the principles appear to apply similarly in the intact and the neurologically impaired system. It is reasonable, then, to hypothesize that the same principles which enhance limb motor learning will also apply to speech motor learning in both healthy and neurologically impaired individuals. AOS is a logical starting point, as it is widely considered a disorder of motor programming. Therefore, clear predictions based on schema theory can be made regarding its response to specific principles of motor learning.

Application of Principles of Motor Learning to Treatment for AOS

Acquired AOS is a motor speech disorder that has been estimated to account for 4% of all acquired neurologic communication disorders (Duffy, 2005). Current research indicates that AOS is a disorder of motor programming (Ballard & Robin, 2007; Clark & Robin, 1998; Deger & Ziegler, 2002; Hageman, Robin, Moon, & Folkins, 1994; Itoh & Sasanuma, 1984; Maas, Robin, Wright, & Ballard, 2008; McNeil, Weismer, Adams, & Mulligan, 1990; for reviews, see Ballard, Granier, & Robin, 2000, and McNeil et al., 1997) that affects programming the kinematic patterns used during speech production (McNeil et al., 1997). Within a motor-programming framework, AOS is a disruption in the ability to select or activate a GMP and/or to select correct parameter values for the execution of movements required for speech production. Motor learning theory, which models the programming of skilled actions, provides an organizing framework that can be applied to the re-learning of speech skills in persons with AOS.

Remediation of AOS has been studied for many years, although long-term retention has rarely been reported in clinical research literature (Wambaugh, Duffy, McNeil, Robin, & Rogers, 2006a, 2006b). A theory-based approach, incorporating principles that enhance learning of motor skills, is lacking in current clinical practice, where anecdotal evidence suggests that speech pathologists tend to use variables that lead to better performance during the therapy session and rarely measure long-term retention and transfer.

Much of the data supporting the application of principles of motor learning to training of speech skills has been derived from healthy populations or limited numbers of speakers with motor speech disorders (e.g.,
Adams & Page, 2000; Ballard, Maas, & Robin, 2007; Knock, Ballard, Robin, & Schmidt, 2000; Maas, Barlow, Robin, & Shapiro, 2002). Clark and Robin (1996) first provided evidence that reduced feedback (KR) frequency facilitates retention of a new oral motor skill in healthy speakers. Steinhauer and Grayhack (2000) subsequently applied the principle of reduced KR frequency to the motor learning of a vowel nasalence task in unimpaired speakers and found an inverse relationship between the frequency of feedback (0%, 50%, or 100%) and measures of performance and learning of the skill. Similarly, Adams and Page (2000) and Adams, Page, and Jog (2002) demonstrated that providing summary feedback (feedback about every trial, presented after a number of intervening trials) after five trials as opposed to providing feedback after every trial enhanced retention of a novel speech skill in normal speakers and speakers with Parkinson’s disease. No studies have yet examined the effect of immediate versus delayed provision of KR feedback on motor speech learning. Overall, the results of early studies suggest that continued work on the influence of principles of motor learning on the speech motor system is warranted.

We designed a series of treatment studies in an effort to understand how feedback (KR) affects acquisition, retention, and transfer of motor speech skills in speakers with AOS. The first two experiments in this series are reported here. We examined the effects of two feedback variables on the treatment of AOS: frequency of feedback and temporal locus of feedback. Although it can be argued that targeting functionally relevant words may be more appropriate in treatment contexts, it was necessary to use nonwords in this study to examine the effects of these principles while avoiding other potentially confounding factors (e.g., concreteness, frequency, familiarity, phonological structure). In other words, this was not a clinical outcomes study; it was a research study to examine the influence of these variables on speech skill learning.

Experiment 1: High- Versus Low-Frequency Feedback

The purpose of Experiment 1 was to examine the effect of frequency of feedback on the learning of speech skills by adults with AOS. Based on evidence from the limb literature (see Schmidt & Lee, 2005, for review), two predictions were made. First, we predicted that high-frequency feedback (HFF) would promote temporary performance enhancement but interfere with long-term retention and transfer of speech skills. The second prediction was that low-frequency feedback (LFF) would best promote the long-term retention of treated speech sounds and facilitate transfer of treated skills to similar but untreated stimuli. These feedback conditions were compared using single-subject design in a common treatment method for AOS.

Method

Participants

Four participants (3 men, 1 woman; M = 70.3 years of age, SD = 3.0 years) with AOS (mean time postonset = 13.3 months; range = 6–20 months; SD = 5.9 months) subsequent to left-hemisphere middle cerebral artery stroke participated in the study. Participants were recruited from the San Diego State University Communications Clinic. Brain scans and detailed lesion information were not available. Three of the participants were right-handed monolingual English speakers. Participant 2 was a left-handed simultaneous bilingual (English and Spanish) speaker with a background in foreign language teaching. He reported that he considered English to be his primary language. Although Participant 1 was 6 months postonset, the possibility of spontaneous recovery did not pose a threat to the validity of his inclusion in the study for two reasons. First, the application of principles of motor learning should impact learning during the subacute stage as well as the chronic stage. Second, the use of the chosen single-subject design (i.e., alternating treatments design; see Experimental Design subsection) allows the separation of the effects of the experimental variables from those related to potential spontaneous recovery. Further, because individuals with AOS receive most of their treatment in the early stages, it is important to study how these variables affect learning at these stages, as well.

We conducted formal testing approximately 2 weeks prior to commencement of the study (see Table 1) and assessed language skills with the Boston Diagnostic Aphasia Examination Battery (BDAE; Goodglass, Kaplan, & Barresi, 2001). Language Competency Indices revealed a wide range of degree of impairment between participants in the language comprehension and expression domains (range = 15th percentile to 81st percentile). To evaluate praxis skills, the Apraxia Battery for Adults-2 (ABA-2; Dabul, 2000) was administered to all participants. Performance on six ABA-2 subtests was analogous to each participant’s BDAE performance, ranging from mild to severe. All participants were diagnosed with AOS by two speech-language pathologists with expertise in motor speech disorders, and all participants produced speech characterized by the following cardinal features of the disorder: increased segmental duration, increased intersegmental duration, errors consisting primarily of distortions, substitutions distorted, consistent error types,
Table 1. Participant characteristics.

<table>
<thead>
<tr>
<th>Participant</th>
<th>Gender</th>
<th>Handedness</th>
<th>Age (yrs)a</th>
<th>Time postonset (mos)b</th>
<th>BDAE percentiles</th>
<th>Language</th>
<th>Diadochokinetic Rate</th>
<th>Word Length</th>
<th>Apraxia</th>
<th>Utterance Time</th>
<th>Polysyllablic Words</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>M</td>
<td>R</td>
<td>74.3</td>
<td>6</td>
<td>15</td>
<td>S</td>
<td>N</td>
<td>N</td>
<td>S</td>
<td>S</td>
<td>N</td>
</tr>
<tr>
<td>P2</td>
<td>M</td>
<td>L</td>
<td>67.3</td>
<td>12</td>
<td>82</td>
<td>Mo</td>
<td>N</td>
<td>N</td>
<td>Mo/S</td>
<td>Mo</td>
<td>S</td>
</tr>
<tr>
<td>P3</td>
<td>M</td>
<td>R</td>
<td>69.0</td>
<td>20</td>
<td>93</td>
<td>Mi</td>
<td>N/Mi</td>
<td>N</td>
<td>Mo</td>
<td>Mo</td>
<td>Mi</td>
</tr>
<tr>
<td>P4</td>
<td>F</td>
<td>R</td>
<td>70.6</td>
<td>15</td>
<td>95</td>
<td>Mi</td>
<td>N/Mi</td>
<td>N</td>
<td>Mi</td>
<td>Mi</td>
<td>Mi</td>
</tr>
</tbody>
</table>

Note. Age and time postonset are given at commencement of Experiment 1. Experiment 2 began 4 months later. BDAE = Boston Diagnostic Aphasia Examination Battery; ABA-2 = Apraxia Battery for Adults-2. yrs = years; mos = months; N = none; Mi = mild; Mo = moderate; S = severe. 

'M' = 70.3, SD = 3.0. 'M' = 13.3, SD = 5.9.
and prosodic anomalies (McNeil et al., 1997; Wambaugh et al., 2006a). These criteria are typical of those used to characterize the perceptual features of AOS in other treatment studies (e.g., Ballard et al., 2007; Wambaugh, Kalinyak-Fliszar, West, & Doyle, 1998; Wambaugh, Martinez, McNeil, & Rogers, 1999; Wambaugh & Nessler, 2004). An oral mechanism examination revealed the probable concomitance of unilateral upper motor neuron dysarthria in Participant 2.

**Experimental Design**

A single-subject alternating treatments design (ATD; McReynolds & Kearns, 1983) was used, with related but untrained behaviors probed throughout the study to assess transfer. An ATD involves administering all treatment conditions (HFF and LFF, in this case) in parallel to all participants. In this way, each participant serves as his/her own control. Each treatment condition must be paired with a different, independent set of behaviors to isolate its effects (e.g., Knock et al., 2000). Treatment condition–behavior set pairing was counterbalanced within participants across two phases of treatment. Counterbalancing of conditions across participants was not possible due to the range of severities and speech impairment profiles represented by the participants.

Because reducing the frequency of feedback is thought to enhance the learning of motor programs as opposed to parameters (Wulf et al., 1993), speech behaviors that were based on different manner classes (e.g., fricatives and plosives; Ballard et al., 2007; Knock et al., 2000; Rubow, Rosenbek, Collins, & Longstreth, 1982) were chosen for Participants 1, 2, and 4 based on their stimulability and profile of impairment (see Table 2). Because differences in manner of production reflect differences in the relative force and timing of muscle contractions, speech behaviors in different manner classes are presumably governed by different GMPs (see Ballard et al., 2007) and, therefore, were considered sufficiently independent to preclude cross-condition contamination. In contrast, place of articulation can be considered a parameter that selects the appropriate muscle groups (or effectors, in schema theory) to execute the program. For Participant 4, stress assignment was also varied, to add complexity (schema theory) to execute the program. For Participant 4, selects the appropriate muscle groups (or effectors, in place of articulation can be considered a parameter that to preclude cross-condition contamination. In contrast, and, therefore, were considered sufficiently independent governed by different GMPs (see Ballard et al., 2007)

Each participant demonstrated stable performance on three baseline probe sessions before treatment began. Each treatment phase was 4 weeks in length (with approximately four treatment sessions per week) with a 4-week maintenance period following (including three to four probes, depending on participant availability). Weekly probe sessions were administered throughout the 16 weeks; these sessions assessed retention of trained behaviors when training conditions were removed and transfer of the trained behaviors to related but untrained targets. In addition, long-term retention data for Phase I were collected on Participants 1 and 4 at eight and seven months, respectively, following the end of Phase I treatment. Participant 2 was unavailable for long-term retention testing, and Participant 3 did not participate because he had suffered a second stroke in the interim.

**Baseline and Probe Testing Procedures**

Baseline and weekly probe sessions consisted of the random elicitation of 10 each of 6 trained nonword behaviors, 12 related but untrained nonword transfer items, and 6 related but untrained real word transfer items, for a total of 240 items for each baseline or probe session. Productions were elicited using orthographic prompts only, and no feedback was given. Rate of stimulus presentation varied as a function of participants’ response times, with the average time between response and presentation of the next stimulus being approximately 2 s. Probes during treatment phases were administered on a day during which no treatment was received.

**Treatment Procedure**

Ninety-minute treatment sessions took place four times per week for a total of 14–16 sessions per participant per treatment phase. The sessions were divided into two periods, with one of the two behavior–condition pairings presented in one period and the other behavior–condition pairing presented in the other period. Order of treatment conditions during each session was counterbalanced across sessions within subject for each treatment phase. Each period began with a pre-practice component, usually 5–15 min in length, involving the use of phonetic placement strategies to elicit at least five correct productions of each of the targets for one behavior set before practice began. The Phonetic Placement Therapy (PPT; Van Riper & Irwin, 1958) involved using orthographic stimuli, pictures, diagrams, verbal descriptions of articulatory features, and/or modeling to shape correct target productions by the participants. For example, if working on /pA/, the clinician might model
Table 2. Design and orthographic stimuli for Experiment 1.

<table>
<thead>
<tr>
<th>Participant</th>
<th>Phase I conditions</th>
<th>Phase I targets</th>
<th>Phase I probes</th>
<th>Phase II conditions</th>
<th>Phase II targets</th>
<th>Phase II probes</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>LFF Fricative</td>
<td>suh</td>
<td>see, us, seam</td>
<td>HFF Fricative</td>
<td>us</td>
<td>suh, ees, cuss</td>
</tr>
<tr>
<td></td>
<td></td>
<td>fuh</td>
<td>fee, uf, fate</td>
<td></td>
<td>uf</td>
<td>fuh, eef, huff</td>
</tr>
<tr>
<td></td>
<td></td>
<td>vuh</td>
<td>vee, uv, vain</td>
<td></td>
<td>uv</td>
<td>vuh, eev, love</td>
</tr>
<tr>
<td></td>
<td>HFF Plosive</td>
<td>tuh</td>
<td>tee, ut, ton</td>
<td></td>
<td></td>
<td>tuh, eet, hut</td>
</tr>
<tr>
<td></td>
<td></td>
<td>puh</td>
<td>pee, up, pat</td>
<td>LFF Plosive</td>
<td></td>
<td>puh, eep, cup</td>
</tr>
<tr>
<td></td>
<td></td>
<td>bluh</td>
<td>bee, ub, beam</td>
<td></td>
<td></td>
<td>bluh, eeb, rub</td>
</tr>
<tr>
<td>P2</td>
<td>LFF Fricative/affricate</td>
<td>chuh-chuh</td>
<td>cheechee, uchuh, chill</td>
<td>HFF Fricative/affricate</td>
<td>uchuh</td>
<td>eechee, chuuchuh, achieve</td>
</tr>
<tr>
<td></td>
<td></td>
<td>thuh-thuh</td>
<td>theethee, uuthuh, thumb</td>
<td></td>
<td>uuthuh</td>
<td>eethee, thuthuh, athena</td>
</tr>
<tr>
<td></td>
<td></td>
<td>zuh-zuh</td>
<td>zeezee, uzhuh, zone</td>
<td></td>
<td>uzhuh</td>
<td>ezeee, zuhzhuh, azores</td>
</tr>
<tr>
<td></td>
<td>HFF L-blends</td>
<td>fluh-fluh</td>
<td>fleehee, ufluh, flame</td>
<td>LFF L-blends</td>
<td></td>
<td>eeflee, fluhfluh, afloaat</td>
</tr>
<tr>
<td></td>
<td></td>
<td>pluh-pluh</td>
<td>pleeplee, upluh, plague</td>
<td></td>
<td></td>
<td>eeplee, pluhpluh, aplotb</td>
</tr>
<tr>
<td></td>
<td></td>
<td>bluh-bluh</td>
<td>bleeblee, ublub, blade</td>
<td></td>
<td></td>
<td>eeblee, blubluh, ablate</td>
</tr>
<tr>
<td>P3</td>
<td>LFF Front-initial</td>
<td>MER-nuh</td>
<td>merNUH, NUHmer, morning</td>
<td>HFF Front-medial</td>
<td>NUH-mer</td>
<td>MER-nuh, nuh-MER, NO-mare</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PER-tuh</td>
<td>perTUH, TUHper, parton</td>
<td></td>
<td></td>
<td>PER-tuh, tuh-PER, temper</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BER-duh</td>
<td>berDUH, DUTHer, bearded</td>
<td></td>
<td></td>
<td>BER-duh, duh-BER, dabber</td>
</tr>
<tr>
<td></td>
<td>HFF Back-initial</td>
<td>HER-nuh</td>
<td>herNUH, NUHher, hornet</td>
<td>LFF Back-medial</td>
<td>NUH-her</td>
<td>HER-nuh, nuh-HER, NO-haer</td>
</tr>
<tr>
<td></td>
<td></td>
<td>KER-tuh</td>
<td>kerTUH, KUTHer, carton</td>
<td></td>
<td></td>
<td>KER-tuh, tuh-KER, tanker</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GER-duh</td>
<td>gerDUH, DUGHer, guarded</td>
<td></td>
<td></td>
<td>GER-duh, duh-GER, dagger</td>
</tr>
<tr>
<td>P4</td>
<td>LFF S-cluster -initial</td>
<td>struh-MUH-nuh</td>
<td>STRUHMuhnuh, mUHNUHSruh, stratify</td>
<td>HFF S-cluster medial</td>
<td>NUH-muh-sruh</td>
<td>STRUHMuhnuh, mUHnuhSTRUH, tapestry</td>
</tr>
<tr>
<td></td>
<td></td>
<td>scruh-MUH-nuh</td>
<td>SCRUMuhnuh, mUHNUHscruh, scrutinize</td>
<td></td>
<td>NUH-muh-scruh</td>
<td>SCRUMuhnuh, mUHnuhSCRUH, redescibe</td>
</tr>
<tr>
<td></td>
<td></td>
<td>spruh-MUH-nuh</td>
<td>SPRUHMuhnuh, mUHNUHspruh, sprucify</td>
<td></td>
<td>NUH-muh-spruh</td>
<td>SPRUHMuhnuh, mUHnuhSPRUH, overspread</td>
</tr>
<tr>
<td></td>
<td>HFF L-blend-initial</td>
<td>fluh-MUH-NUH</td>
<td>FLUHMuhnuh, mUHNUHFLUH, flocculate</td>
<td>LFF L-blend medial</td>
<td>nuh-MUH-fluh</td>
<td>fluhMuhnuh, nuhMNUHFLUH, megaflp</td>
</tr>
<tr>
<td></td>
<td></td>
<td>bluh-MUH-NUH</td>
<td>BLUHMuhnuh, mUHNUHBLUH, bloviate</td>
<td></td>
<td>nuh-MUH-bluh</td>
<td>bluhMuhnuh, nuhMNUHBLUH, notably</td>
</tr>
<tr>
<td></td>
<td></td>
<td>gluh-MUH-NUH</td>
<td>GLUHMuhnuh, mUHNUHGLUH, glomarize</td>
<td></td>
<td>nuh-MUH-gluh</td>
<td>gluhMuhnuh, nuhMNUHGLUH, polyglot</td>
</tr>
</tbody>
</table>

Note. Capital letters indicate stress. LFF = low-frequency feedback; HFF = high-frequency feedback.
the target, instruct the participant to look at the clinician and listen carefully to the sound, show “puh” on a card orthographically, and say (for example) “for this sound, start with the lips pressed together.” Feedback involving knowledge of results (KR; whether the sounds were correct or incorrect) and knowledge of performance (KP; how the sounds were produced; e.g., “your lips were apart”) were given during PPT pre-practice. Consistent with previous studies examining motor learning in AOS (e.g., Knock et al., 2000), PPT was selected because it is part of almost all treatments for AOS across a range of severities. Moreover, PPT results in clear acquisition effects under controlled experimental conditions (see Wambaugh & Doyle, 1994, for review). This pre-practice was followed by a practice component in which 30 productions of each target in a set were elicited in random order with orthographic prompts only, for a total of 90 productions. General KR feedback (i.e., “correct” or “incorrect”) was provided on 60% (LFF) or 100% (HFF) of productions. During both pre-practice and practice, the feedback interval and postfeedback interval each approximated 2 s. For trials in which feedback was not given, intertrial intervals were approximately 2 s. For the LFF condition, feedback schedules were constructed beforehand and were used online by the clinician to ensure reliability of the independent measure. Three different LFF schedules were used in order to avoid the same trials receiving feedback during each session.

Stimuli and Materials

Due to the wide range of speech and language skills observed in these participants, each received individually tailored stimulus sets appropriate to his or her profile of impairment and stimulability (see Table 2). For all participants, three nonword syllables or syllable sequences were trained under each feedback condition (e.g., “suh” or “MER-nuh”; see Table 2). Transfer was tested to untrained stimuli of similar complexity (e.g., “see” and “us” or “mer-NUH” or “NUH-mer,” respectively) and to more complex real words related to the trained behaviors (e.g., “seam” or “morning,” respectively). Word frequency was balanced across both sets of complex real words in each treatment phase (Francis & Kuçera, 1982) for Participants 1, 2, and 4. Complex real-word stimuli for Participant 4 were verbs (for homogeneity and conformity with carrier phrase restrictions) with very low frequency (1 or less; Francis & Kuçera, 1982); however, because a three-syllable English verb beginning with the /spr/ cluster was not available, a pseudoword was constructed and defined for the participant (“sprucify: to spruce or clean up the appearance of something”). Because the participant was unfamiliar with most of the other real words, this decision seemed justified. For Participant 3, it was not possible to balance the sets of real-word stimuli for word frequency due to the need to control for other potentially more influential factors such as stimulability. However, there was no difference in baseline performance on these sets of real-word stimuli in Phase I, and the real-word stimuli chosen for Phase II were all highly infrequent (Francis & Kuçera, 1982).

The use of nonwords is motivated by several points. First, the consensus is that AOS is a disorder of speech motor programming, and nonwords (here, syllables licensed in English) also involve the relevant processes, from phonological encoding through motor programming through articulation. Also, to remove possible confounds arising at prior (semantic and/or lexical) stages of processing as a result of concomitant aphasia, we decided to isolate the relevant processes via the use of nonwords. In addition, there is evidence to support the use of nonwords in treatment of AOS, in that generalization has been observed to real words (e.g., Maas et al., 2002; Schneider & Frens, 2005). However, it is possible that using real words would lead to greater learning. To our knowledge, no studies are available that have directly compared the efficacy of using real words versus nonwords in this population and, as such, it remains an open question as to which is better. For example, Tjaden (2000) included both nonwords and words; however, these were not counterbalanced (nonwords always preceded words), and this study was not designed to compare the two.

All stimuli were presented to participants orthographically. Prior to the beginning of the study, all participants were screened for their ability to read the stimuli by demonstrating reliability in matching examiner-produced tokens to the appropriate written stimulus in a field of six. For Participant 1, who was severely impaired, place-of-articulation photographs were also used during the practice component of the treatment sessions and during a “modified” probe (see Modified probe subsection on the next page). For Participant 2, target behaviors were embedded in easy or difficult carrier phrases (e.g., “It’s a ___ a day” or “The hyperthyroidism will ___ the ray”; see Appendix) in Phases I and II of the study, respectively. Targets for Participant 4 were embedded in difficult carrier phrases in both phases (see Appendix).

Scoring and Reliability

All productions were scored online as correct or incorrect by the examiner and up to two other clinicians for purposes of determining interrater reliability. Only the examiner’s judgment counted for purposes of providing feedback. Responses were considered correct if all phonemes were accurately produced (without distortion),
if the appropriate stress pattern was applied (in multisyllabic productions), and if the rate of the production was deemed “normal.” Error-free carrier phrase production was not required for a “correct” score on a target item. For Participant 2, coarticulation across syllables was not required for a score of “correct” during Phase I, as he was encouraged to think of the targets as reduplicated consonant-initial nonwords for the early goal of achieving correct articulation of these consonants in the initial position of syllables (for which he was most stimulable). The second goal, of producing these consonants in the medial position of a two-syllable nonword without segmentation, was encouraged during the second phase, when coarticulation across syllables was required for a score of “correct.” Participant responses were also recorded via digital audiotape (DAT) recorder for future analyses.

Interrater reliability was measured for a randomly selected 25% (Participant 1: 26%, Participant 2: 29%, Participant 3: 22%, Participant 4: 24%) of treatment and probe sessions. Point-to-point agreement on “correct” and “incorrect” binary judgments was 95.3% (fricatives: 95.5%, plosives: 95.1%) for Participant 1, 83.6% (fricatives/affricates: 85.5%, L-blends: 81.7%) for Participant 2, 86.2% (front phonemes: 85.8%, back phonemes: 86.8%) for Participant 3, and 88.9% (S-clusters: 91.9%, L-blends 85.8%) for Participant 4.

**Results and Discussion**

**Participant 1, Phase I**

**Baseline.** Participant 1 (P1) demonstrated stable baseline performance on fricative and plosive VC behaviors (e.g., “fuh” and “tuh”) before Phase I treatment began. Accuracy ranged from 0% to 5% for items to be trained (see Figure 1, Panel A, Sessions 1–3) and from 0% to 8% for untrained transfer items (see Figure 1, Panels B, C, and D, Sessions 1–3), with no substantial differences across conditions.

**Acquisition.** Improvements in performance during treatment sessions (see Figure 1, Panel A, filled shapes) were observed for trained behaviors for both treatment conditions (LFF and HFF). The rates of improvement and accuracy levels achieved did not differ for LFF and HFF items, which reached 69% and 67%, respectively.

**Retention and transfer.** P1’s probe performance in Phase I revealed no changes in accuracy of production for retention of trained items (see Figure 1, Panel A, unfilled shapes) and transfer to untrained items (see Figure 1, Panels B, C, and D, Phase I) in either feedback condition (see subsequent **Retention and transfer** subsection on this page for note on word transfer items in Panel D).

**Participant 1, Phase II**

**Baseline.** P1 demonstrated stable baseline performance on fricative and plosive VC behaviors (e.g., “ut” and “ut”) prior to treatment. Accuracy rates were similar across conditions and ranged from 0% to 5% for items to be trained (see Figure 1, Panel B, Sessions 1–21).

**Acquisition.** P1 acquired the HFF-fricative behaviors at a faster rate and to a slightly higher level (80%) than the LFF-plosive behaviors (58%; see Figure 1, Panel B, filled shapes), supporting our first prediction that HFF would enhance initial skill acquisition. Interestingly, after only one session of Phase II treatment, P1 improved from 0% to 50% accuracy on the HFF-fricative behaviors (see Figure 1, Panel B, Sessions 24–25). This could be due to a very early benefit of HFF for acquisition or, alternatively, a longer-term effect of LFF-fricative training carried over from Phase I. HFF throughout Phase II treatment likely compounded this effect.

**Retention and transfer.** Probe performance in Phase II suggests that P1 was unable to transfer trained behaviors to untrained items in either condition (see Figure 1, Panels C–D, Phase II). Small fluctuations in performance on word transfer items (see Figure 1, Panel D) are not considered reflective of feedback condition influence because their baseline was not perfectly stable and also because they consistently favor plosive items across both phases (i.e., both feedback conditions). However, examination of the retention of trained VC targets in regular probe sessions (see Figure 1, Panel B, Phase II, unfilled shapes) suggests that beginning in Phase II treatment and continuing throughout maintenance, this participant was able to demonstrate learning in the difficult context of the probe session format (consisting of the random elicitation of 24 different targets 10 times each), and his performance consistently favored the LFF condition (LFF range: 0%–43%; HFF range: 0%–23%).

**Modified probe.** Due to P1’s extreme difficulty producing trained behaviors in the large probe stimulus set, and his inconsistent and perseverative error patterns, it was felt that his retention of trained speech skills was not reflected by probe performance. Therefore, a limited set of trained items only, elicited five times each with photographic support, was presented every week after Phase I treatment as a “modified probe” (see Figure 1, Panel E). As with the regular probes, no feedback was given during this modified probe. Due to P1 traveling, only two modified probes were administered during the maintenance stage of Phase I. The first maintenance probe was administered the day after P1 returned and may reflect his reported jetlag. However, a modified probe 1 week later (4 weeks after Phase I treatment, Session 23) illustrated strong retention of both sets of
Figure 1. P1 performance during baseline, treatment, and maintenance segments of the two phases of Experiment 1. Filled shapes in Panels A and B represent performance during high-frequency and low-frequency feedback (HFF and LFF, respectively) practice trials in treatment sessions. Unfilled shapes in Panels A and B represent baseline and retention of trained behaviors in probe sessions. Panels C and D represent transfer to related but untrained stimuli. (Note that different transfer stimuli were probed in each of the two phases.) Panel E depicts retention of trained behaviors in modified probe sessions. Session LT represents a long-term retention probe that occurred 8 months following the end of Phase I treatment.
trained behaviors but no evident difference in retention between the two sets of behaviors under different feedback conditions in Phase I (LFF = 74%; HFF = 72%). When long-term retention was assessed with the modified probe task 8 months after the termination of Phase I treatment (with Phase II and Experiment 2 intervening), a small difference in retention (LFF = 54%; HFF = 38%) between the two conditions was detected (see Figure 1, Panel E, Session LT). Speech targets that were trained under reduced feedback conditions were retained better at 8 months post-treatment. The other therapies that the patient received in the intervening months were not likely contributors to this effect, as they did not involve production of Phase I targets and there is no evidence that this speaker was able to transfer skills between related items. Time constraints did not permit the collection of long-term retention data 8 months after Phase II.

Consistent with the postulated carry-over effect from LFF conditions in Phase I (see Phase II Acquisition section) is P1’s performance on the first session of the modified probe testing in Phase II (see Figure 1, Panel E, Session 28), in which LFF-plosives (formerly HFF) were 0% accurate, and HFF-fricatives (formerly LFF) were almost 70% accurate. The carry-over hypothesis for these results seems tenable, especially after examining the time-course of modified probe performance throughout Phase II. By the end of Phase II maintenance, retention of HFF behaviors (as measured via modified probes) was roughly unchanged from this first probe session, whereas retention of LFF behaviors had increased. This result, along with the outcome of the long-term retention probe in Phase I (Session LT) and the maintenance of performance in Phase II (see Figure 1, Panel B, unfilled symbols), supports the second prediction that LFF leads to enhanced learning.

**Participant 2, Phase I**

*Baseline.* Participant 2 (P2) demonstrated stable baseline performance on L-blend and fricative/affricate C(c)VC(c)V behaviors (e.g., “pluhpluh” and “zuhzuh”) before Phase I treatment began (see Figure 2, Panel A, Sessions 1–3), with correct responses ranging from 0% to 10% for items to be trained (see Figure 2, Panel A, Sessions 1–3) and from 0% to 17% for untrained transfer items (see Panels B, C, and D, Sessions 1–3), with no substantial differences across conditions.

*Acquisition.* Contrary to the first prediction, P2’s acquisition performance on trained behaviors was better for LFF-fricative/affricate items (range: 5%–97%) than for HFF-L-blend items (range: 2%–71%; see Figure 2, Panel A, filled shapes).

*Retention and transfer.* Retention (see Figure 2, Panel A, unfilled shapes) and transfer (see Figure 2, Panels B–D, Phase I) were greater and more stable for items related to trained LFF-fricative/affricate behaviors than for items related to trained HFF-L-blend behaviors, appearing to lend support to the second experimental prediction. Accuracy of responses to retention probes of trained stimuli during treatment and maintenance periods of Phase I ranged from 3% to 97% for LFF-fricative/affricate behaviors and from 0% to 43% for HFF-L-blend behaviors. Transfer to untrained productions involving different vowels (LFF range = 2%–80%; HFF range = 0%–40%; see Panel C) and to untrained words (LFF range = 42%–83%; HFF range = 6%–63%; see Panel D) demonstrated a similar pattern of results. A large degree of transfer was observed to the VCV stimuli that were going to be used as treatment targets for Phase II (see Figure 2, Panel B, Phase I). Because this improvement continued to grow and fluctuate drastically throughout Phase I Maintenance, an additional probe was administered before Phase II Baseline began. These items were returned to near-zero accuracy levels by increasing their difficulty for Phase II (see earlier sections on Stimuli and Materials and Scoring and Reliability).

**Participant 2, Phase II**

*Baseline.* After implementing the above-mentioned difficulty manipulations, P2 demonstrated stable baseline performance on L-blend and fricative/affricate VC(c)V behaviors (e.g., “uhpluh” and “uzuhuh”) before Phase II treatment began (see Figure 2, Panel B, Sessions 28–30). Response accuracy ranged from 0% to 3% for items to be trained in the LFF condition and from 0% to 7% in the HFF condition. Baseline accuracy for untrained transfer behaviors ranged from 0% to 10% for LFF items and from 0% to 20% for HFF items.

*Acquisition.* Performance on targets given HFF (now fricatives/affricates) was better during acquisition than those given LFF (now L-blends; see Figure 2, Panel B, filled shapes). Performance accuracy ranged from 3% to 87% for HFF targets and from 0% to 61% for LFF targets.

*Retention and transfer.* Retention (see Figure 2, Panel B, Phase II, unfilled shapes) and transfer to untrained items (see Figure 2, Panels C–D, Phase II) were also slightly better in the HFF condition (HFF retention range = 50%–87%; LFF retention range = 23%–33%). This pattern of results across both phases of the study, wherein fricative/affricate behaviors fairly consistently outperformed L-blend behaviors regardless of feedback conditions, suggests a probable stimulus effect. It is likely that the additional phonological and motoric complexity of L-blends as compared with fricative/affricates influenced accuracy enough to conceal any underlying effect of feedback frequency in this participant.
Figure 2. P2 performance during baseline, treatment, and maintenance segments of the two phases of Experiment 1. Filled shapes in Panels A and B represent performance during HFF and LFF practice trials in treatment sessions. Unfilled shapes in Panels A and B represent baseline and retention of trained behaviors in probe sessions. Panels C and D represent transfer to related but untrained stimuli. (Note that different transfer stimuli were probed in each of the two phases.) More difficult carrier phrases were introduced at the onset of Phase II. fric./aff. = fricative/affricate.
Participant 3, Phase I

Baseline. Participant 3 (P3) demonstrated stable baseline performance with zero accuracy on back-initial and front-initial behaviors (e.g., “KERtuh” and “MERnuh”) before Phase I treatment began (see Figure 3, Panel A, Sessions 1–3).

Acquisition. P3’s performance on both sets of behaviors improved with treatment (see Figure 3, Panel A, filled shapes), but contrary to our first prediction, HFF did not appear to promote substantially better performance in acquisition of speech skills than LFF did (HFF range = 0%–82%; LFF range = 0%–76%).

Retention and transfer. Despite the absence of marked acquisition differences, retention (see Figure 3, Panel A, unfilled shapes) and transfer (see Figure 3, Panels C–D, Phase I) of trained behaviors were enhanced in the LFF condition, as predicted. Ranges of accuracy for trained items (LFF = 0%–53%; HFF = 0%–25%), untrained stress-transfer items (LFF = 0%–63%; HFF = 0%–27%), and untrained word-transfer items (LFF = 0%–63%; HFF = 0%–33%) revealed obvious overall enhanced learning in the LFF condition.

Participant 3, Phase II

Baseline. Stable zero-accuracy baseline performance on back-medial and front-medial behaviors (e.g., “TUHker” and “NUHner”) was demonstrated before Phase II treatment began (see Figure 3, Panel B, Sessions 1–26).

Acquisition. P3’s performance of both behaviors improved markedly with Phase II treatment (see Figure 3, Panel B, filled shapes). However, as in Phase I, HFF did not enhance performance during acquisition of speech skills relative to LFF, contrary to our first prediction (HFF range = 0%–79%; LFF range = 6%–89%).

Retention and transfer. Although overall enhanced learning under conditions of LFF was more obvious in Phase I, Phase II did provide evidence of earlier retention in probe sessions for trained targets given LFF (see Figure 3, Panel B, Sessions 31 and 35). Also, despite the retention of trained HFF behaviors reaching a higher peak right after treatment (HFF = 70%; LFF = 33%; see Figure 3, Panel B, Session 47), it soon began to decline (Sessions 48–49; final probe HFF = 27%), whereas behaviors trained in the LFF condition in Phase II were retained at a level similar to that achieved during the treatment phase (final probe LFF = 43%). An apparent contradiction to this pattern was the real-word transfer set in Phase II, which appeared to be enhanced under HFF conditions (see Figure 3, Panel D, Phase II). These data, however, should be interpreted with caution because the baseline level was higher for real words related to the HFF than the LFF condition (see Figure 3, Panel D, Session 27). When pre-treatment and post-maintenance endpoints for these Phase II real-word data sets were compared (LFF = 0%, 10%; HFF = 26%, 33%), the LFF condition indeed was associated with as much change as the HFF condition.

Participant 4, Phase I

Baseline. Participant 4 (P4) demonstrated stable zero-accuracy baseline performance on S-cluster and L-blend behaviors (e.g., “struh-MUH-nuh” and “fluh-muh-NUH”) before Phase I treatment began (see Figure 4, Panel A, Sessions 1–3). Baseline accuracy for untrained order and stress transfer behaviors (see Figure 4, Panels C–D, Sessions 1–3) was also zero, whereas untrained word transfer items (see Panel E) ranged from 7% to 27% for HFF items and from 7% to 8% for LFF items.

Acquisition. In the treatment setting, P4 acquired L-blend behaviors given HFF at a faster rate and to a higher level than S-cluster behaviors given LFF, in accord with our first prediction (see Figure 4, Panel A, filled shapes). Accuracy rates during acquisition ranged from 3% to 92% for HFF items and from 0% to 49% for LFF items.

Retention and transfer. Contrary to our second prediction, retention (see Figure 4, Panel A, unfilled shapes) and transfer (see Figure 4, Panels C–E, Phase I) were stronger for the L-blend behaviors given HFF than for the S-cluster behaviors given LFF. It is noteworthy, however, that after the termination of treatment, retention of behaviors given HFF steadily declined from 93% to 63% across the four maintenance probes, whereas retention of behaviors given LFF remained stable or increased slightly (from 33% to 47%), relative to the level achieved during treatment (see Figure 4, Panel A, Sessions 23–26). A long-term follow-up probe session indicated that P4 retained much of the improvement in trained skills (HFF = 60%; LFF = 40%) 7 months after the termination of Phase I treatment (see Figure 4, Panel A, Session LT). Performance levels on transfer to untrained stress and word transfer behaviors (see Figure 4, Panels C–E, Session LT) were also maintained to varying degrees, with HFF L-blend skills continuing to surpass LFF S-cluster skills. The only other treatment that this participant received in the intervening months was Experiment 1 Phase II treatment, which did not involve production of Phase I targets. Time constraints did not permit the collection of long-term retention data 7 months after Phase II.

Participant 4, Phase II

Baseline. A stable baseline with zero accuracy was demonstrated on syllable-order reversed S-cluster and
Figure 3. P3 performance during baseline, treatment, and maintenance segments of the two phases of Experiment 1. Filled shapes in Panels A and B represent performance during HFF and LFF practice trials in treatment sessions. Unfilled shapes in Panels A and B represent baseline and retention of trained behaviors in probe sessions. Panels C and D represent transfer to related but untrained stimuli. (Note that different transfer stimuli were probed in each of the two phases.)
Figure 4. P4 performance during baseline, treatment, and maintenance segments of the two phases of Experiment 1. Filled shapes in Panels A and B represent performance during HFF and LFF practice trials in treatment sessions. Unfilled shapes in Panels A and B represent baseline and retention of trained behaviors in probe sessions. Panels C, D, and E represent transfer to related but untrained stimuli. (Note that different transfer stimuli were probed in each of the two phases.) Session LT represents a long-term retention probe that occurred 7 months following the end of Phase I treatment.
L-blend behaviors (e.g., “NUH-muh-STRUH” and “nuh-MUH-fluh”) before Phase II treatment commenced (see Figure 4, Panel B, Sessions 27–29). Untrained transfer items were also stable at or near zero accuracy (see Figure 4, Panels C–E, Sessions 27–29).

**Acquisition.** Although P4 made gains in both behavior sets during Phase II treatment, HFF applied to the training of S-cluster behaviors did not result in better performance, contrary to prediction. A greater level of acquisition was achieved for L-blend behaviors given LFF (see Figure 4, Panel B, filled shapes), with correct responses ranging from 1% to 56% for HFF items and from 0% to 83% for LFF items.

**Retention and transfer.** After the termination of treatment, probe performance of many trained and untrained items continued to improve throughout the maintenance phase, but a difference between HFF and LFF conditions was not evident in retention of trained targets (see Figure 4, Panel B, unfilled shapes) or transfer to untrained real words (see Figure 4, Panel E, Phase II). Transfer to different syllable order (HFF range = 0%–25%; LFF range = 0%–38%) and stress pattern (HFF range = 3%–8%; LFF range = 7%–57%) was better for LFF-L-blend behaviors (see Figure 4, Panels C–D, Phase II), apparently supportive of our second hypothesis. However, as was the case with the other mildly impaired participant (P2), the results for P4 are also suggestive of a stimulus effect. In general, across both phases of the study, P4 acquired (see Figure 4, Panels A–B), retained (see Figure 4, Panels A–B), and transferred (see Figure 4, Panels C–E) best the L-blend as opposed to the S-cluster behaviors, regardless of feedback condition, in accord with hypothesized differences in phonologic and motoric complexity.

---

**Experiment 2: Immediate Versus Delayed Feedback**

Experiment 2 examined the effect of feedback timing on speech skill learning in adults with AOS by comparing the effects of instantaneous feedback versus delayed feedback on acquisition and learning. On the basis of recent evidence in the limb literature (e.g., Anderson et al., 2005), two predictions were made. First, we predicted that immediate feedback (IF) would promote temporary performance enhancement but would interfere with retention and transfer of speech skills. Second, we predicted that delayed feedback (DF) would best promote retention of treated speech sounds and transfer of treated skills to similar but untreated stimuli. These feedback conditions were also compared using single-subject design in a common treatment method for AOS.

### Method

#### Participants

Two participants who had completed participation in Experiment 1 (P1, age 74.7, and P2, age 67.7) with chronic AOS (time postonset 10 and 16 months, respectively) subsequently to left-hemisphere stroke participated in the second study, which commenced approximately 1 week later. The third participant from Experiment 1 began Experiment 2 but withdrew when he suffered another stroke. The fourth participant from Experiment 1 did not participate in Experiment 2 because it was not possible to construct a new set of stimuli that were challenging enough given her mild impairments yet could also be reliably judged online by the examiner. Formal testing was conducted approximately 4 months prior to commencement of the study. Participant details and test results were reported for Experiment 1 (see Table 1).

#### Experimental Design

The design of Experiment 2—comparing immediate and delayed feedback—was identical to that of Experiment 1. Stimuli were again chosen based on the assumption that different production manners represented different motor programs (Ballard et al., 2007) and therefore could presumably be trained concurrently under different feedback conditions without introducing cross-condition contamination.

#### Baseline and Probe Testing Procedures

Baseline and weekly probe sessions were similar to those described in Experiment 1 except that because of the differences in participant ability and stimulus set makeup, composition of probes differed between participants and between elicitation of 180 (P1) or 240 (P2) items for each baseline or probe session. Productions were elicited via orthographic prompts only.

#### Treatment Procedure

Treatment procedures were similar to those of Experiment 1. Instead of being given verbally (as in Experiment 1), general knowledge-of-results feedback was provided visually by showing the participant a red or green signal (for “incorrect” or “correct”) and was provided on 100% of trials either immediately (IF; instantaneously [within 1 s]) upon completion of the trial or after a 5-s delay (DF) following the participant’s production. The postfeedback delay interval was a constant 5 s in each of these conditions. Because the intertrial interval includes both of these segments, it was also necessarily different between conditions (5 s for IF, 10 s for DF).
Stimuli and Materials

Participants in Experiment 2 continued to present with different levels of speech and language skills and, therefore, stimuli were individually tailored for appropriateness to each participant’s profile of impairment and stimulability (see Table 3). Participants were given training on production of three nonword syllables or syllable sequences under each feedback condition (e.g., “shee” or “VAYmuhnay”; see Table 3). Transfer to similar but untrained behaviors (e.g., “eesh” or “muhNAYvay” and “FAYmuhnay,” respectively) and to real words related to the trained behaviors (e.g., “sheep” or “vaporize,” respectively) was assessed in weekly probe sessions. It was not possible to control for word frequency of these real-word stimuli due to the need to control for other potentially more influential factors such as a participant’s baseline stimulability for these words (as in P1) or the unavailability of English single words related closely enough to the trained nonwords (as in P2). Real-word stimuli used for P2 were either of very low frequency (4 or less; Francis & Kučera, 1982) or were “composite” words made of real words and/or morphemes and defined for the participant (e.g., “misinvoke”). P2 was encouraged to think of and produce “composite” word stimuli made of more than one real word (e.g., “visacard”) as single words. As before, speech behaviors were elicited by orthographic printing of stimuli on cards.

As in Experiment 1, due to P1’s difficulty in demonstrating retention of familiar trained items during standard probe sessions, a “modified probe” was administered for each treatment set during the regular probe session each week after the termination of Phase I treatment.

Table 3. Design and orthographic stimuli for Experiment 2.

<table>
<thead>
<tr>
<th>Variable</th>
<th>P1</th>
<th>P2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IF</td>
<td>DF</td>
</tr>
<tr>
<td>Phase I conditions</td>
<td>fricative</td>
<td>plosive</td>
</tr>
<tr>
<td>Phase I targets</td>
<td>shay</td>
<td>kay</td>
</tr>
<tr>
<td></td>
<td>shee</td>
<td>kee</td>
</tr>
<tr>
<td></td>
<td>shoo</td>
<td>kao</td>
</tr>
<tr>
<td>Phase I probes</td>
<td>Aysh</td>
<td>ayk</td>
</tr>
<tr>
<td></td>
<td>eesh</td>
<td>eek</td>
</tr>
<tr>
<td></td>
<td>oosh</td>
<td>ook</td>
</tr>
<tr>
<td></td>
<td>shame</td>
<td>came</td>
</tr>
<tr>
<td></td>
<td>sheep</td>
<td>keep</td>
</tr>
<tr>
<td></td>
<td>shut</td>
<td>cut</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Phase II conditions</th>
<th>plosive</th>
<th>fricative</th>
<th>R-blend</th>
<th>fricative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase II targets</td>
<td>ayk</td>
<td>aysh</td>
<td>muhNAYbray</td>
<td>muhNAYvay</td>
</tr>
<tr>
<td></td>
<td>eek</td>
<td>eesh</td>
<td>muhNEEbre</td>
<td>muhNEEvee</td>
</tr>
<tr>
<td></td>
<td>ook</td>
<td>oosh</td>
<td>muhNObro</td>
<td>muhNOvo</td>
</tr>
<tr>
<td>Phase II probes</td>
<td>kay</td>
<td>shay</td>
<td>BRAYmuhnay</td>
<td>VAYmuhnay</td>
</tr>
<tr>
<td></td>
<td>kee</td>
<td>shee</td>
<td>BREEmuhnne</td>
<td>VEEmuhnee</td>
</tr>
<tr>
<td></td>
<td>koo</td>
<td>shoo</td>
<td>BROmuhno</td>
<td>VOmuhno</td>
</tr>
<tr>
<td></td>
<td>back</td>
<td>lash</td>
<td>muhNAYpray</td>
<td>muhNAYfay</td>
</tr>
<tr>
<td></td>
<td>walk</td>
<td>wash</td>
<td>muhNEEpree</td>
<td>muhNEEfee</td>
</tr>
<tr>
<td></td>
<td>look</td>
<td>bush</td>
<td>muhNOpro</td>
<td>muhNOfo</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>celebrate</td>
<td>salivate</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>gentle-breeze</td>
<td>old-TVs</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>go-for-broke</td>
<td>misinvoke</td>
</tr>
</tbody>
</table>

Note. For P2 stimuli, capital letters indicate stress. IF = immediate feedback; DF = delayed feedback.
(starting with Session 23 and continuing throughout Phase II). This modified probe essentially segmented the probe task into two smaller retention tests (one for each stimulus set) consisting of trained items only (elicited randomly 10 times each) followed by a transfer test consisting of the random elicitation of the remaining 12 untrained items (elicited randomly 10 times each). As in the standard probe protocol, feedback was withheld during modified probes.

Scoring and Reliability

All productions were again scored online as correct or incorrect following the same guidelines as established and described in Experiment 1. Interrater reliability was calculated on a randomly selected 13% (P1: 12%, P2: 14%) of treatment and probe sessions. Point-to-point agreement for “correct” and “incorrect” perceptual judgments was 94.4% (plosives: 94.8%, fricatives: 94.1%) for P1 and 97.8% (R-blends: 98.1%, fricatives: 97.5%) for P2.

Results and Discussion

Participant 1, Phase I

Baseline. P1 demonstrated stable baseline performance on trained fricative and plosive CV targets (e.g., “shoo” and “koo”) before Phase I treatment began (see Figure 5, Panel A, Sessions 1–3). Accuracy ranged from 0% to 3% for items to be trained and from 0% to 7% for

Figure 5. P1 performance during baseline, treatment, and maintenance segments of the two phases of Experiment 2. Filled shapes in Panels A and B represent performance during immediate and delayed feedback (IF and DF, respectively) practice trials in treatment sessions. Unfilled shapes in Panels A and B represent baseline and retention of trained behaviors in probe sessions. All probes after the withdrawal of Phase I treatment were elicited in the “modified” format (see text discussion). Panel C represents transfer to related but untrained stimuli. (Note that different transfer stimuli were probed in each of the two phases.)
untrained transfer items (see Figure 5, Panels B–C, Sessions 1–3), with no substantial differences across conditions.

**Acquisition.** During Phase I treatment, acquisition of trained behaviors was better under conditions of DF than IF (see Figure 5, Panel A, filled shapes), contrary to the first prediction. Correct responses ranged from 20% to 77% for DF items and from 2% to 57% for IF items.

**Retention and transfer.** P1’s probe performance in Phase I revealed enhanced retention of trained skills in the DF condition (DF range = 0%–90%; IF range = 0%–67%; see Figure 5, Panel A, unfilled shapes), in accord with our second prediction, but revealed mixed results for untrained transfer probes (see Figure 5, Panels B–C, Phase I). Specifically, effects of DF treatment transferred to the opposite phoneme order (VC; DF range = 0%–23%; IF range = 0%; see Figure 5, Panel B, Phase I), whereas the effects of IF treatment showed greater transfer to the real-word stimuli (DF range = 0%–27%; IF range = 3%–50%; see Figure 5, Panel C, Phase I). P1’s performance on real-word probes might be explained by differences in his personal experience with the specific words (e.g., he associated one of the fricative words with an important political figure at the time and relished telling that for every opportunity to practice it), so it is perhaps more telling that for “phoneme order” nonword probes (see Figure 5, Panel B, Phase I), P1 was never able to transfer to this behavior in the IF condition but was able to do so (although not to a great extent) in the DF condition.

**Participant 1, Phase II**

**Baseline.** A stable baseline was demonstrated on fricative and plosive VC targets (e.g., “oosh” and “ook”) prior to the commencement of the study (accuracy range: 0%–13%), although slight instability was induced before Phase II treatment began when improved performance on plosive behaviors transferred from Phase I treatment (see Figure 5, Panel B, Session 12). The resultant discrepancy between plosive and fricative skills was small (M = 8.5%) and relatively unchanged before Phase II treatment began. Baseline accuracy for untrained word transfer items was 0% for DF behaviors and from 3% to 7% for IF behaviors (see Figure 5, Panel C, Sessions 26–27).

**Acquisition.** Although P1 achieved 94% accuracy on both behavior sets by the end of Phase II treatment, his overall acquisition performance levels were better for the targets given IF (range: 53%–94%) than those given DF (range: 14%–94%), as predicted (see Figure 5, Panel B, filled shapes).

**Retention and transfer.** Retention of the trained skills in the probe situation (see Figure 5, Panel B, unfilled shapes) paralleled this acquisition pattern (i.e., there was stronger retention of skills given IF [range: 57%–87%] than those given DF [range: 33%–57%]), contrary to the second prediction. Transfer of treatment effects to related items was highly variable in Phase II, and a consistent effect of feedback delay was not evident. Given that P1’s performance largely favored plosives in acquisition, transfer, and retention in both phases, differences may be attributable to a stimulus effect and may not reflect a response to feedback timing conditions.

**Participant 2, Phase I**

**Baseline.** P2 demonstrated stable zero-accuracy baseline performance on R-blend and fricative target behaviors (e.g., “BRAY-muh-nay” and “VAY-muh-nay”) before Phase I treatment began (see Figure 6, Panel A, Sessions 1–3). Baseline accuracy levels for untrained transfer items (see Figure 6, Panels B–D, Sessions 1–3) were at or near zero.

**Acquisition.** As predicted, skills given IF were performed better (range: 19%–100%) during acquisition than those trained under conditions of DF (range: 0%–74%) in Phase I (see Figure 6, Panel A, filled shapes).

**Retention and transfer:** Probe data suggest that retention (see Figure 6, Panel A, unfilled shapes) and transfer (see Figure 6, Panel D, Phase I) were enhanced by provision of DF. Although retention of treated skills in the IF condition initially led retention of treated skills in the DF condition (see Figure 6, Panel A, Sessions 22–25), 1 week into the maintenance phase (Session 24), retention of IF targets began to decline (from 93% to 43%), whereas retention of DF targets remained stable or showed slight improvement throughout the maintenance phase (from 30% to 47%) and the second phase of the study (to 80%). Inspection of untreated word transfer items (see Figure 6, Panel D, Phase I) in Phase I yields similar interpretation: Transfer of IF treatment effects to untrained items appeared temporarily enhanced during the treatment phase (i.e., Sessions 17 and 22), but some of this improvement fluctuated during the maintenance phase (range: 13%–23%), whereas transfer to untrained words related to the DF condition continued to improve throughout the maintenance phase (range: 10%–37%; see Figure 6, Panel D, Sessions 23–26). Transfer to untrained voiceless phonemes (see Figure 6, Panel C, Phase I) is discussed in the subsequent section.

**Participant 2, Phase II**

**Baseline.** A stable or slightly declining baseline on syllable-order reversed R-blend and fricative behaviors (e.g., “muh-NAY-bray” and “muh-NAY-vay”) was demonstrated before Phase II treatment began (see Figure 6, Panel B, Sessions 27–28). Accuracy rates for training items averaged 12% for IF behaviors and 2% for DF behaviors. Baseline levels for untrained transfer items were at or near zero (see Figure 6, Panels C–D, Sessions 27–28).
**Figure 6.** P2 performance during baseline, treatment, and maintenance segments of the two phases of Experiment 2. Filled shapes in Panels A and B represent performance during IF and DF practice trials in treatment sessions. Unfilled shapes in Panels A and B represent baseline and retention of trained behaviors in probe sessions. Panels C and D represent transfer to related but untrained stimuli. (Note that different transfer stimuli were probed in each of the two phases.)
Acquisition. In Phase II, targets subjected to IF were again performed better (range: 3%–77%) during acquisition than those given DF (range: 1%–63%), even though stimulus–condition pairings were reversed from Phase I (see Figure 6, Panel B, filled shapes), supporting our first prediction.

Retention and transfer. Enhanced learning under conditions of feedback delay was evident in most trained and untrained items. On initial inspection (data not shown), P2 did not evidence retention or transfer (no improvement above baseline) of the trained behaviors for either condition in the probe setting, wherein these 6 items were mixed with 18 other transfer items (elicited 10 times each). A gradual change in response quality, however, was noticed by the experimenter, and thus a formal error analysis was subsequently conducted. Many of P2’s productions contained multiple errors, including incorrect stress, distorted phonemes, and reduced articulatory speed. As P2 acquired the new skills in treatment sessions, the number of different errors in each production decreased, and more productions, although still scored as “incorrect,” were characterized by a speed-only error. Therefore, to capture this improvement on retention and transfer items, the data were re-analyzed with a more lenient criterion; the speed-only errors were also considered “correct.” Using the more lenient criterion, Figure 6 illustrates better overall retention (see Panel B, unfilled shapes) and transfer (see Panels C–D, Phase II) in the DF condition, consistent with the interpretation suggested by the more stringent criterion for accuracy. Retention accuracy ranged from 17% to 70% in the DF condition, and from 13% to 57% in the IF condition (see Panel B). Transfer to untrained words ranged from 0% to 83% in the DF condition, and from 0% to 57% in the IF condition (see Panel D). One exception to this pattern may be observed with inspection of the untrained voicing transfer items, which appeared enhanced consistently with the fricative targets in both phases (see Figure 6, Panel C). A possible cause for this is that these probes involved voicing of a previously trained voiced consonant, and this skill may be more complex in the context of an R-blend than in the fricative (singleton) condition, thus obscuring potential effects of the feedback timing manipulation on this type of transfer probe.

General Discussion

Taken as a whole, the data reported here (see Table 4 for summary of results) provide some support for the a priori hypotheses but also raise important challenges for a simple interpretation relative to feedback. The results of 2 of the 4 participants studied in Experiment 1 (P1 and P3) support the prediction that reducing the frequency of feedback enhances learning (retention and transfer). Of these 2 participants, the first prediction (i.e., that HFF would be associated with temporary performance enhancement during initial speech skill acquisition) was only borne out in 1 participant and in only one phase of his treatment. More important, however, is that despite these inconsistent acquisition patterns in P1 and P3, LFF was associated with enhanced long-term retention and/or transfer of trained speech skills in both participants. Furthermore, comparison of P1’s Phase I and Phase II target acquisition data suggests that LFF during training of a particular behavior (in this case, fricatives in CV syllables in Phase I) may also carry over to boost future acquisition of that behavior in a different context (e.g., VC syllables in Phase II). The differential effects of HFF and LFF in this participant, as well as his demonstration of stable baselines, indicate that these changes cannot be solely attributed to spontaneous recovery during the subacute stage. The effect of feedback frequency in the 2 remaining participants (P2 and P4) was masked by the probable influence of stimulus complexity. That is, enhanced performance across both phases of the study was associated with phonologically and motorically simpler targets (see Barlow & Gierut, 1999)—namely, singletons versus L-blends for P2 and two-element blends versus three-element clusters for P4.

The results of Experiment 2 offer qualified support for the prediction that IP enhances the initial acquisition of speech skills. P1’s performance in both phases of the study revealed a pattern of learning (i.e., retention and transfer) that roughly paralleled acquisition under each of the two feedback conditions. That is, enhanced learning was observed in whichever condition was also associated with enhanced acquisition. Moreover, the enhanced performance was consistently associated with plosives relative to fricatives, suggesting that it may have been
driven by a stimulus complexity effect that was not overridden by an effect of feedback delay. P2 demonstrated the predicted pattern of IF associated with enhanced acquisition and DF associated with enhanced learning (i.e., retention and transfer). This did not reflect a stimulus effect, as the same pattern was observed in both phases. That is, under all stimulus–condition pairings, IF and DF were consistently associated with enhanced acquisition and learning, respectively.

Motor learning theory asserts that the availability of outcome information is a crucial factor for learning and that the frequency and temporal locus of external feedback may determine the availability of the outcome information to the learner and its effect on long-term learning. Frequent and immediate provision of augmented feedback may be detrimental to long-term learning because both may interfere with information processing of internal outcome information. According to the “guidance hypothesis,” HFF is thought to enhance performance during training by providing increased guidance and heightening energy and motivation (Lee et al., 1990; Salmoni et al., 1984). However, these performance effects are temporary and are dependent on the continued delivery of the feedback. Presenting KR feedback with lower frequency is hypothesized to enhance long-term learning by facilitating the development of self-evaluation and error-detection skills that the learner can apply to situations in which external feedback is not available (Bruechert, Lai, & Shea, 2003). Similarly, imposing a feedback delay should promote the development of these skills by allowing the learner sufficient time to process and build relationships (schemas) based on the four kinds of information discussed previously. Indeed, participants in the Anderson et al. (2005) novel aiming task study reported using a greater variety of intrinsic feedback sources under conditions of feedback delay than with immediate or no feedback. Alternatively, it has been suggested that reducing the frequency of augmented feedback may serve to relieve the performer of a state of internal attentional focus induced by constant feedback, and to allow the performer to assume an external attentional focus guided by automatic motor control processes (Wulf, McConnel, Gärtnér, & Schwarz, 2002). This theory argues that an external attentional focus benefits both acquisition performance and learning. Although the current studies were not aimed at testing this theory, it may provide an explanation (that the guidance hypothesis cannot) for the lack of demonstrable acquisition benefits of IF or HFF for some participants in the present experiments.

For impaired learners, particularly those with cognitive–linguistic or attentional deficits, a hypothetical “optimal window” for temporal locus of feedback can be conceived in which the learner has sufficient time to process this information but not so much time that the activation of the information cannot be sustained. Immediately and/or consistently filling the postmovement interval with external outcome information may initially facilitate acquisition by reducing the cognitive demands of activating and maintaining in working memory the internal outcome information, but instantaneous and frequent provision of external feedback may also block the learner from learning how to activate and maintain this information when external outcome information is withdrawn.

The data from the current studies contribute new evidence to the growing body of literature supporting the application of principles of motor learning to the motor speech system, but these data also raise many important questions. One question pertains to why the feedback manipulations were strong factors affecting acquisition and learning for some individuals but not for others. Furthermore, when comparing across studies, different participants appear to have benefited from different feedback manipulations. Curiously, no single participant demonstrated a benefit from both reduced frequency and delayed feedback. The answer may involve individual variation in the hypothesized “optimal window” for receiving external feedback and may be based on ability to meet the attentional demands for activation and maintenance of various information types (cf. Li & Wright, 2000). Another potential explanation might involve the particular manifestation of AOS in individuals. Clark and Robin (1998) have suggested that AOS may reflect disruption of generalized motor programs in some patients but a disruption of parameterization (of intact motor programs) in others. If these categories indeed represent two different etiological subtypes of AOS, we may expect to find individual differences in the effectiveness of feedback manipulations, depending on the specific nature of the underlying impairment in an individual. However, our dependent measure of perceptual accuracy does not allow us to differentiate these potentially different subtypes. Finally, it is unclear which aspects of the stimuli used here are best conceived of as governed by GMPs versus parameters. Because it has been suggested that these feedback manipulations operate differently on GMPs and parameters (e.g., Wulf et al., 1993), potential modification of the predictions for learning of these skills should be addressed in future studies.

Another important issue in deciphering the effects of particular motor learning conditions involves the potentially additive or interactive nature of multiple conditions of practice. For example, Wulf and colleagues (2002) reported an interaction between relative feedback frequency and type of feedback. Specifically, they demonstrated that reduced feedback frequency was only influential in the learning of sport skills under internal-focus (as opposed to external-focus) feedback conditions. In the current studies, all feedback given in Experiment 1
(feedback frequency) was provided immediately after production, and in Experiment 2 (feedback delay), the feedback frequency was a constant 100%. In addition to the feedback frequency and delay manipulations, other principles of motor learning were inherent to the design of this study. For example, trials of all behaviors in a given set (e.g., three fricative targets) were practiced in random order, and random practice has been shown to enhance retention and transfer of motor skills relative to blocked practice and may facilitate learning by increasing the difficulty of the learning environment and by approximating a natural context (Knock et al., 2000). In addition, as noted earlier, some participants demonstrated a learning pattern of consistent higher accuracy for less-complex skills (e.g., L-blends vs. S-clusters), regardless of the feedback condition with which they were paired, suggesting that stimulus complexity may override the effects of the feedback variables tested here. Finally, and perhaps most importantly, a very large number of repetitions of each behavior was elicited over the course of the treatment phases in each of the studies. Perhaps the most influential of the training conditions, amount of practice, may have obscured the potentially more subtle effects of the feedback manipulations. These potential interactions between multiple conditions of practice and feedback would best be addressed in the future by a group treatment design.

A further concern is that most of the work that has been done on motor learning for limb movements has used kinematic measurements, not perceptual accuracy judgments, to assess learning of movement skills. Perceptual accuracy judgments were used in this study because of their clinical relevance and ecological validity. However, using perceptual accuracy measures to make inferences about processes occurring at the kinematic level may not be appropriate. It is likely that perceptual judgments, although valuable for functional measurement in clinical outcome studies, may not correlate with kinematic values and that gross accuracy judgments do not capture more subtle changes that occur at the kinematic level as a result of training under different conditions (Ballard & Robin, 2002; Ballard et al., 2007).

Finally, it is clear that treatments for AOS (those reported in this article and those in the literature) are effective for some participants to varying degrees and are not effective for other participants, although it is important to note that we observed either benefits to learning of reduced or delayed feedback or no difference between conditions, whereas there were no instances in which 100% or immediate feedback unambiguously enhanced learning. Understanding why particular experimental manipulations work or do not work for a given individual will ultimately require examination of treatment outcomes for a given therapy approach with regard to pre-treatment variables that may help define for whom a given treatment is effective. Such pre-treatment variables that will be important to consider in future research include severity of concomitant cognitive, linguistic, and motoric symptoms, neurological factors such as lesion size, and demographic factors such as time postonset and education level. For example, although participants in the present study were screened informally for their ability to read the orthographic stimuli, it is possible that the presence of concomitant aphasia interacted with the motor learning studied here.

In conclusion, although treatment protocols for AOS abound in the clinical practice of speech-language pathology, few studies have been undertaken to investigate the effects of theory-driven practice factors in the remediation of this disorder in a systematic fashion. The data from the current studies, as part of a programmatic line of research, contribute new evidence to the growing body of literature supporting the application of principles of motor learning to the motor speech system and support the larger notion that the speech motor learning and control system shares properties with limb motor learning and control. In addition, they endorse the principle well-documented in the motor learning literature: that acquisition performance does not necessarily predict true learning. Independent of feedback manipulations, these data also contribute to the growing body of literature supporting the efficacy of treatment for chronic AOS.

**Acknowledgments**

This work was supported, in part, by National Institute on Deafness and Other Communication Disorders (NIDCD) Pre-Doctoral Training Grant DC007361, a VA Pre-Doctoral Research Fellowship (both awarded to the first author), San Diego State University Foundation Grant GIA 0304 (awarded to the second author), and NIDCD Grant DC005698 (awarded to the fourth author).

**References**


Received February 28, 2006
Revision received July 12, 2007
Accepted November 29, 2007
DOI: 10.1044/1092-4388(2008/06-0042)

Contact author: Shannon N. Austermann Hula, who is now with the VA Pittsburgh Healthcare System, Geriatric Research, Education and Clinical Center, Pittsburgh, PA 15206. E-mail: shannon.hula@va.gov.

Donald A. Robin is now affiliated with the Human Performance Division, Research Imaging Center, departments of Neurology and Radiology and the Program in Biomedical Engineering, The University of Texas Health Science Center at San Antonio, and with the Honors College and College of Liberal and Fine Arts, The University of Texas, San Antonio.

Edwin Maas is now affiliated with the Department of Speech, Language, and Hearing Sciences at the University of Arizona.

Kirrie J. Ballard is now affiliated with the Faculty of Health Sciences, University of Sydney, Australia.
**Appendix.** Carrier phrases used in Experiment 1.

<table>
<thead>
<tr>
<th>Easy</th>
<th>“He likes to ___ each day.”</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>“It’s a ___ a day.”</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Difficult</th>
<th>“The thermoelectricity may ___ the town.”</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>“The municipality may ___ the pen.”</td>
</tr>
<tr>
<td></td>
<td>“The hospitalization may ___ the week.”</td>
</tr>
<tr>
<td></td>
<td>“The hypothyroidism may ___ the ray.”</td>
</tr>
<tr>
<td></td>
<td>“The misinterpretation may ___ the net.”</td>
</tr>
</tbody>
</table>

Austermann Hula et al.: Feedback in AOS Treatment