Ultrasound visual feedback treatment and practice variability for residual speech sound errors

Jonathan L. Preston,
Haskins Laboratories, New Haven, CT & Southern Connecticut State University, New Haven, CT

Patricia McCabe,
The University of Sydney, Sydney & Murdoch Children's Research Institute, Melbourne, Australia

Ahmed Rivera-Campos,
University of Cincinnati, Cincinnati, OH

Jessica L. Whittle,
Haskins Laboratories, New Haven, CT

Erik Landry, and
Southern Connecticut State University, New Haven, CT & Haskins Laboratories, New Haven, CT

Edwin Maas
University of Arizona, Tuscon, AZ

Abstract

Purpose—The goals were to (1) test the efficacy of a motor-learning based treatment that includes ultrasound visual feedback for individuals with residual speech sound errors, and (2) explore whether the addition of prosodic cueing facilitates speech sound learning.

Method—A multiple baseline single subject design was used, replicated across 8 participants. For each participant, one sound context was treated with ultrasound plus prosodic cueing for 7 sessions, and another sound context was treated with ultrasound but without prosodic cueing for 7 sessions. Sessions included ultrasound visual feedback as well as non-ultrasound treatment. Word-level probes assessing untreated words were used to evaluate retention and generalization.

Results—For most participants, increases in accuracy of target sound contexts at the word level were observed with the treatment program regardless of whether prosodic cueing was included. Generalization between onset singletons and clusters was observed, as well as generalization to sentence-level accuracy. There was evidence of retention during post-treatment probes, including at a two-month follow-up.

Conclusions—A motor-based treatment program that includes ultrasound visual feedback can facilitate learning of speech sounds in individuals with residual speech sound errors.

Residual speech sound errors (RSSEs) are a subtype of speech sound disorder in which sounds remain in error beyond the typical age of acquisition (i.e., beyond 8–9 years of age)

Corresponding Author: Dr. Jonathan Preston, Haskins Laboratories 300 George St, Ste 900 New Haven, CT 06511, preston@haskins.yale.edu 203-865-6163 x247.

(Preston & Edwards, 2007; Shriberg, Austin, Lewis, McSweeny, & Wilson, 1997). Individuals with RSSEs are beyond the developmental window in which speech sound normalization is likely to occur spontaneously. Many individuals with RSSEs produce substitutions or distortions of later developing sound such as /ɹ, l, s, z, ʃ, tʃ, θ/, sometimes despite years of therapy (Flipsen, 2003; Gruber, 1999; Irwin, Huskey, Knight, & Oltman, 1974; Sax, 1972). Some individuals with RSSEs do not respond well to traditional treatment methods (e.g., McAllister Byun & Hitchcock, 2012). Therefore, there is a critical need to develop and evaluate efficacious alternate treatments.

Research evaluating treatments for RSSEs is surprisingly limited, despite the fact that a large percentage of speech-language pathologists provide therapy for these individuals (American Speech-Language-Hearing Association, 2010). Traditional approaches to therapy are still frequently used, including elicitation procedures such as shaping and phonetic placement cueing with drill and repetition (Secord, Boyce, Donohue, Fox, & Shine, 2007; Shriberg, 1975, 1980; Shriberg & Kwiatkowski, 1982; Van Riper & Erickson, 1996). Other approaches to remediating RSSEs involve visual feedback of acoustic cues (McAllister Byun & Hitchcock, 2012; Shuster, Ruscello, & Smith, 1992) or tongue-palate contact using electropalatography (Dagenais, 1995; Dent, Gibbon, & Hardcastle, 1995; Gibbon & Hardcastle, 1987). Thus, instrumental methods may be implemented to provide visual feedback on articulatory performance. The present study focuses on ultrasound as an instrumental method for providing visual feedback of the tongue.

A common framework for characterizing treatment research has been a multi-phase progression (Fey & Finestack, 2009; Robey, 2004). In Phase I, the purpose of the research is to establish the initial concept as viable; Phase I studies typically have low levels of experimental control and are often reported as case reports or small *n* pre-post group studies. Treatment studies typically progress through Phase II (operationalizing the treatment procedures and outcomes, demonstration of effect), Phase III (efficacy, typically in randomized controlled trials), Phase IV (effectiveness in field trials in clinical settings) and ideally Phase V (cost-benefit analysis). The present study represents Phase II research investigating the use of real-time ultrasound visual feedback of tongue shapes/movements with individuals with RSSEs. Ultrasound imaging allows the client and clinician to observe tongue position and shape, to directly cue changes in tongue position or shape, and to evaluate whether the client has achieved the intended changes. To date, Phase I evidence suggests that ultrasound visual feedback may have beneficial effects for individuals with RSSEs. Ultrasound feedback has been explored with children, adolescents and adults with idiopathic RSSEs (Adler-Bock, Bernhardt, Gick, & Bacsfalvi, 2007; Lipetz & Bernhardt, 2013; Modha, Bernhardt, Church, & Bacsfalvi, 2008; Shawker & Sonies, 1985), speech sound errors secondary to hearing impairment (Bacsfalvi, 2010; Bacsfalvi & Bernhardt, 2011), and speech errors following stroke (Preston & Leaman, 2014). Additionally, recent Phase II evidence of treatment effect in improving articulatory accuracy has been shown with children with speech sound errors associated with Childhood Apraxia of Speech (Preston, Brick, & Landi, 2013). Therefore, it is appropriate to conduct Phase II type research examining ultrasound visual feedback in treatment of children with idiopathic RSSEs.

Motor Learning

Children with RSSEs may have trouble establishing an appropriate motor plan for sounds (such as achieving a particular tongue shape), or they may have trouble achieving proper timing of movements (such as coordinating an appropriate tongue shape with movements that happen before or after the sound) or finally, they may have learned an incorrect plan, which once fixed in place is difficult to change. Thus, the present investigation uses elements of schema based motor learning theory (Maas et al., 2008; Schmidt & Lee, 2011) to guide the structure of the treatment sessions. Motor learning principles have been well described elsewhere (see Maas et al., 2008, for a review); however, we briefly summarize the elements as they pertain to the present investigation.

To establish a new motor skill, differential treatment has to apply to skill acquisition as compared with skill learning. We define skill acquisition as a temporary change in performance while we define skill learning as a relatively permanent change in behavior (e.g., Schmidt & Lee, 2011). In order to establish the new skill an individual must both acquire and then learn the skill. Research has shown in a range of motor tasks that skill acquisition can be enhanced by high frequency feedback on how the task is performed (knowledge of performance: KP) as well as judgments of correct or incorrect (knowledge of results: KR). In the current study, KP feedback is provided by the ultrasound image which is visible to and used by both the clinician and the participant during each relevant block of treatment. KP verbal feedback is also provided by the clinician on a reducing scale over successful trials.

To retain and generalize the new motor skill, feedback should be reduced in frequency, randomized as to which trial receives feedback, and changed in nature from KP (with or without KR) to KR alone. When learning a motor skill, it has been proposed that a general motor plan (GMP) is developed for that movement (Schmidt & Lee, 2011). GMPs are specific to a given movement but can be applied variably within more complex movements through changing their parameters (e.g., speed and amplitude/force of movement). In speech, GMPs may be hypothesized to include stored instructions for phonemes, gestures, syllables or words, while speech related parameters may include variations in timing and magnitude related to rate, pitch, intensity, and force. Research on non-speech motor learning has suggested that once a GMP has been established, transitioning from constant to variable practice (i.e., a range of parameters) leads to a more readily retrieved motor plan which can be flexibly applied to a range of uses (Hall & Magill, 1995; Wulf & Schmidt, 1997). In speech-related research, Ballard, Maas, and Robin (2007) observed positive treatment effects when including variable practice for adults with acquired apraxia of speech. Studies that have directly compared constant vs. variable practice in speech treatment have found mixed results; Wong, Ma and Yiu (2011) reported no advantage for practice variability among individuals with hyperfunctional voice problems but Adams and Page (2000) observed an advantage for variable practice in motor speech learning in healthy adults. However, there is no systematic exploration of this principle of practice variability among speech motor learning in individuals with RSSEs. In the current study, to change the parameters of speech we used prosodic variability. That is, individuals plan and execute speech by integrating a variety of segmental and suprasegmental commands, and individuals

learning new GMPs for sounds in error must learn to achieve target lingual movements while varying additional respiratory and phonatory demands. By requiring participants to produce the motor plan in a range of prosodic environments we hypothesized that we were varying the movement parameters and thus facilitating learning and generalization. Prosodic variation was included in conjunction with an ultrasound treatment protocol with children with CAS (Preston et al., 2013), and it is often included as one element in other motor-based treatments as well (Strand & Debertine, 2000; Strand, Stoeckel, & Baas, 2006).

For optimal learning to occur, treatment sessions should be structured such that practice is neither too easy nor too hard. This concept has been referred to as the "challenge point" (Guadagnoli & Lee, 2004; Rvachew & Brosseau-Lapré, 2012). Therefore, there should be constant modifications of task difficulty to advance from simpler to more complex items when the individual is ready (e.g., from syllables to monosyllabic words, multisyllabic words, and phrases). Attention should also be paid to the amount of information provided to the client, beginning with high frequency feedback with primarily KP and transitioning to lower frequency feedback that is primarily KR. For learning to occur, modifications should be made to both the task and to the information available, and the learner is expected to make mistakes. Thus, 100% accuracy within a session is not necessarily the goal (as this would mean that the task is too simple or too much support has been provided), and a high degree of accuracy within the session is not the long-term goal. Learning is evident when a skill is retained over time and when high performance can be achieved in untrained contexts without feedback.

In addition to evaluating motor learning through retention, we also considered generalization effects. Generalization can be gleaned from evidence of transfer to untreated words, or to untreated phonetic contexts. Recent studies have reported no generalization effects from onset to coda word positions when training /ɹ/ (McAllister Byun & Hitchcock, 2012; Preston et al., 2013), and this may be due to the distinct differences in timing and magnitude of syllable-initial and syllable-final productions of /ɹ/ (Campbell, Gick, Wilson, & Vatikiotis-Bateson, 2010); however, Preston et al. (2013) observed generalization from trained to untrained /ɹ/ singletons and clusters in the onset position of words using ultrasound visual feedback (see also A. L. Williams, 1991 for evidence of such generalization). The present study further explores these generalization effects in onset position as evidence of motor learning in individuals with RSSEs.

Purpose and hypotheses

The purpose of the study was to develop and to test the effects of a motor-learning based treatment for individuals with RSSEs using ultrasound visual feedback. We hypothesized that individuals with RSSEs could show improved accuracy with an ultrasound visual feedback program that is based on motor learning principles. We also hypothesized that the use of increased practice variability through a condition which included variable prosodic cueing would result in larger effect sizes for generalization items compared with a condition that did not include such prosodic cues during practice of words and phrases. Finally, we hypothesize that there would be generalization from trained singletons to clusters in the onset position.

Method

Participants

Participants were at least 10 years old and had residual speech sound errors. They were recruited by referrals from local Connecticut speech-language pathologists. To be eligible for the study, participants had to be native English speakers and had to pass a hearing and vision screening. Participants had to score below a standard score of 75 on the Goldman-Fristoe Test of Articulation-2 (Goldman & Fristoe, 2000) and below 20% accurate on at least two probes developed to provide deeper assessments of late-developing sounds (see Target Selection below). All participants had errors on /ɹ/ and, in some cases, other sounds as well. Additionally, participants had normal nonverbal cognition, as defined by Perceptual Reasoning Index scores above 80 on the Wechsler Abbreviated Scales of Intelligence-2 (Wechsler, 2011). Those with diagnosed developmental disabilities such as Autism or Cerebral Palsy were excluded from the study. Language skills were required to be broadly within the normal range as defined by scores above 80 on the Peabody Picture Vocabulary Test-4 (Dunn & Dunn, 2007) and scaled scores above 7 on the Recalling Sentences and Formulated Sentences subtests of the Clinical Evaluation of Language Fundamentals-4 (Semel, Wiig, & Secord, 2003). Additionally, to characterize language performance, the Expressive Vocabulary Test-2 (Williams, 2007) and the Comprehensive Test of Phonological Processing-2 (Wagner, Torgesen, Rashotte, & Pearson, 2013) were also administered. Participant characteristics are described in Table 1. Ten children were referred; one failed to meet the inclusion criteria as her speech was not sufficiently impaired, therefore nine participants commenced treatment. One participant was unable to adhere to the study protocol and discontinued the study after six sessions and is not reported further. The study therefore included eight participants, six males and two females. All but one (participant #85) had received previous therapy. As much of the treatment took place during school break in the summer, only one participant (#74) was receiving treatment by another speech-language pathologist at the time of the study, and that SLP agreed to treat other communication domains (morphosyntax and literacy) while the study was in progress.

In addition, a battery was administered to evaluate speech motor skills. This included (a) a motor speech protocol that involved sustained productions of consonants and vowels as well as rapid productions of syllables and syllable sequences (Rvachew, Hodge, & Ohberg, 2005), (b) a multisyllabic word imitation task (Preston & Edwards, 2007), (c) a token-to-token inconsistency test, with 10 polysyllabic words repeated 10 times each (adapted from Dodd, Hua, Crosbie, Holm, & Ozanne, 2002), (d) an emphatic stress task, in which participants repeated three four-word sentences with varied stress (cf. Shriberg et al., 2010), and (e) a conversational speech sample. From these tasks, three speech-language pathologists familiar with CAS judged the participants on a three point scale as 0= no apparent signs of CAS, 1=possible CAS, and 2=likely CAS. All participants reported here received 0 scores by all three listeners, indicating agreement that these participants did not currently exhibit symptoms of this speech sound disorder subtype. That is, there were no obvious deficits with sound sequencing or transitioning, sound consistency, syllable segregation, or prosody. Quantitatively, seven of the participants completed the motor

speech protocol (Rvachew et al., 2005) and all seven received dysarthria scores of 0 and apraxia scores of 0, further ruling out motor speech impairments.[1]

Target selection

Probes were developed to assess accuracy of late-developing sounds in various contexts. A context was defined by word position and sound, such as initial /ɹ/ or final /s/. Probes were made for initial and final /s/, /z/, /θ/, /ʃ/, /tʃ/, /ɹ/, as well as initial /ɹ/, /s/, and /l/ clusters. Each probe consisted of 25 words, and these were administered by having the child read the words one at a time with no cueing. During the first evaluation session, the SLP identified sounds in error during conversation and the GFTA-2 and selected appropriate sounds to be probed. If the participant achieved less than 20% correct (fewer than 5/25), the context was available for treatment or for monitoring. These probes were re-administered before treatment to obtain 3–4 baseline data points. Following the onset of treatment, probes were administered every-other session (at the beginning of sessions before treatment began, alternating between different probes each session). Two sound contexts that were below 20% accuracy were selected for treatment. Each context was treated for 7 sessions, with the order randomly assigned. When possible, untreated sounds that were phonetically unrelated to the treated sound contexts were probed to add additional experimental control. When word-initial singletons or clusters were treated, probes were administered to evaluate generalization from singletons to clusters (or vice versa) in word-initial position.

Design

The study used a multiple baseline across behaviors single subject experimental design, replicated across eight participants. Probes were administered three to four times before initiating treatment to evaluate stability of baselines. Treatment began on one sound context in one of the two treatment conditions (Prosodic Cueing or No Prosodic Cueing, see below), with the order of conditions counterbalanced (half of the participants began in Prosodic Cueing) and the two treated sound contexts randomly assigned to condition. Thus, for each participant, one sound context was treated in the first phase of the study (first 7 sessions) while other contexts were untreated. Between the first and second treatment phases, three to five probes were recorded. Then, the second treatment phase commenced by targeting a different sound context in the other condition (Prosodic Cueing or No Prosodic Cueing). Three to five post-treatment probes were also administered. Treatment effects were observed if the treated sound contexts showed improvement over pre-treatment levels. Experimental control was confirmed by monitoring the untreated sound contexts which were phonetically dissimilar to treated sounds (and therefore not expected to show generalization from the treatment). Retention was evaluated by repeating the probes at a two month follow-up session.

Generalization was observed by evaluating accuracy on word-level probes (25 words that were not treated) as well as accuracy in sentences. Each participant was recorded completing the same 15 sentence imitation task, which was administered before treatment and during a

---

[1]Participant 83 was unable to complete the motor speech protocol without coughing due to illness; thus, a valid score was not obtained.

session within one week post-treatment. The sentences included a total of 45 /s/ and 53 /ɹ/ in all word positions. Two research assistants (graduate students with training in phonetics and speech sound disorders) who were blind to pre- or post-treatment status scored accuracy of /s/ and /ɹ/ productions from audio recordings presented in random order.

Treatment

The ultrasound feedback protocol used was modified from prior work with children with CAS and an adult with acquired apraxia of speech (Preston et al., 2013; Preston & Leaman, 2014). The same certified SLP was responsible for providing treatment to all participants, although a second certified SLP who was highly familiar with the protocol provided treatment on 13% of sessions when the primary clinician was unavailable. During the course of the treatment, no homework activities were given to the participants. Each participant was treated on two sound contexts which were assigned to the conditions outlined below.

Condition differences—For four participants, the first context was treated in Prosodic Cueing and the second context was treated with No Prosodic Cueing; for the other four participants, the order of conditions was reversed. Prosodic Cueing simply required selecting a new cue card after every three trials. These cards showed one of a question mark (?), an exclamation point (!) or a period (.) and were designed to cue the child to produce the target words or phrases with prosodic variations. Regardless of whether the examiner perceived the intended prosody, feedback was only given on segmental accuracy. Therefore, the nature of the practice varied between conditions but the nature of feedback was identical in the two conditions.

Condition similarities—Treatment sessions were 60 mins long, with the first 6–8 mins for administering probes, followed by four 13 min Periods (a timer was used to ensure adherence to the 13 min time limits). Periods A and C included ultrasound visual feedback, whereas Periods B and D included no ultrasound visual feedback. These Periods could be characterized primarily as drill or drill-play (Shriberg & Kwiatkowski, 1982). To maintain motivation, Periods B and D optionally included taking a turn in a brief motivational activity (e.g., Connect-4, dice games, a basketball toss, computer games such as Angry Birds®, etc.) after each block of six trials at a syllable, word, or phrase. Some participants chose not to include these activities and elected to spend all of the Periods in drill-related practice.

For each context that was treated, two variants were targeted. For consonant singleton targets, these variants included different vowels before or after the consonant; for example, for the context initial /ɹ/, two variants that were targeted included /ɹe/ and /ɹo/. When clusters were targeted, the two variants included two different clusters such as /fɹ/ and /gɹ/.

Period A always began with an Elicitation period. Elicitation was designed to provide instruction to the child about how to produce the target sound context. The Elicitation period was structured in such a way that the clinician could address each child's specific needs, as the cues given were dependent upon the sound context and the child's specific error type. During the Elicitation period, both a coronal and a sagittal view could be used with the ultrasound, and multiple cues were given related to tongue shape and position. The goal of the Elicitation period was for the child to achieve six correct productions of each variant.

For example, for the context initial /ɹ/, six correct productions of both /ɹe/ and /ɹo/ had to be achieved before moving to the Structured Practice and Feedback period (see below). The primary strategies used in the Elicitation period were phonetic placement cues (e.g., "move your tongue further back"), shaping (e.g., shaping /a/ or /l/ into [ɹ], or shaping /ʃ/ into [s], etc.), along with visual feedback with the ultrasound. A transparent sheet was placed over the computer screen. Marks were put on the sheet to provide "targets" for the participant to "hit" with the tongue or to trace a shape that the participant was instructed to try to match (cf. Preston et al., 2013). For participants who were not stimulable for correct productions, an entire session could be spent in Elicitation. Alternately, for participants who were readily stimulable, the Elicitation period could be completed in less than a minute.

Structured Practice and Feedback

Once the child met the criteria for the Elicitation period during that session (six correct productions of each variant), Structured Practice and Feedback period was used for the remainder of the session. The schedule of the Structured Practice and Feedback was a modification of a practice schedule developed in a previous study (Preston & Leaman, 2014). Identical practice and feedback schedules were used in the Structured Practice and Feedback period regardless of whether the ultrasound visual feedback was available (Periods A and C) or whether there was no ultrasound visual feedback (Periods B and D).

During the Structured Practice and Feedback period, there were five levels of increasing complexity: syllables (which included only an onset or a coda), monosyllabic words (which included an onset and a coda), multisyllabic words, set phrases, and phrases containing a cloze structure. Practice occurred in blocks of six trials (practice attempts). To begin the Structured Practice and Feedback, six trials occurred at the syllable level addressing one variant of the sound context. For example, if initial /ɹ/ was the sound context and /ɹe/ and /ɹo/ were the two treated variants, the client began with six practice trials at /ɹe/. If the child did not achieve the step-up criterion of at least five of six correct productions of the variant, then s/he made six practice trials at the other variant (e.g., /ɹo/). If the step-up was not met, the child returned to six trials at the first variant (e.g., /ɹe/). However, if the step-up criterion was met, the child progressed to the second level, monosyllabic words. This included a block of six trials at a single word containing the variant that the child just successfully produced. Hence, if the child successfully produced at least five of six /ɹe/ syllables, s/he then made six trials at a word such as "rain" because this word contained the target syllable. As the client continued to meet the step-up criterion, practice advanced to multisyllabic words, phrases, and cloze phrases. However, failure to meet Level Criterion at any level (monosyllabic words, multisyllabic words, or phrases) meant that the child returned to the syllable level of the other variant (e.g., returning to /ɹo/ syllables). This meant that the child was able to quickly progress through increasing levels of complexity from syllables through phrases only when successful; however, the child could also spend the entire session practicing at the syllable level if s/he was not meeting the step-up criterion. The Appendix provides a flow diagram outlining an example of this progression.

At the cloze phrase level, participants were required to generate their own phrase that either began or ended with the target (e.g., for participants working on /aɹ/, phrases such as

"_____ car" might be elicited, with the participant expected to generate phrases such as "blue car" and "fast car"). For participants who met the step-up criterion at the cloze phrase level for the same word in two blocks of a session, the words/phrases deemed "successful" were removed from the practice list and replaced with equivalents in subsequent sessions. Additionally, if a participant successfully met the step-up criterion for cloze phrases of three different words during two blocks in a single session, then the duration of the ultrasound Period was reduced by half the following session (e.g., from two 13-min Periods to one 13 min Period, or from one 13 min Period to one 6 min Period, or from 6 min to no ultrasound).

Across the five levels of complexity, different feedback schedules were provided. As the child progressed to higher levels, the aims were to (1) gradually reduce the frequency of feedback, and (2) transition from providing primarily Knowledge of Performance (KP) providing primarily Knowledge of Results (KR). Table 2 summarizes the feedback schedule. The data sheets used to guide the practice schedule and feedback are available as Supplemental Material, and the full manual used to guide the treatment is available for free by contacting the first author.

Use of ultrasound to provide KP

A Seemore PI 7.5 MHz probe was placed under the chin during Periods A and C and images were visible by both the child and the clinician. Verbal feedback specific to the child's error was provided. For example, errors on /ɹ/ might involve a high tongue dorsum and low tongue blade, and the KR+KP on these productions (using a sagittal view) included statements such as "No, the dorsum was too high" or "Good, the tip was up." Lateralized distortions of sibilants were cued (using a coronal view) by describing a groove in the middle of the tongue and elevation of the sides of the tongue. Feedback included statements such as "No, you need a deeper groove" or "Yes, you got the sides of the tongue high." Dentalized productions of /s/ were cued (using a sagittal view) by pointing out the constriction location and encouraging a more posterior constriction (Lipetz & Bernhardt, 2013). A supplemental video is available online demonstrating the use of ultrasound.

Dose

In accordance with principles of motor learning, children were expected to complete a substantial number of practice attempts (cf. Edeal & Gildersleeve-Neumann, 2011). During treatment, the average number of practice attempts per session was 210 (SD 44; n.b., this does not include trials from the Elicitation period, as only practice trials in the Structured Practice and Feedback period were counted). The number of trials was evenly distributed between ultrasound and non-ultrasound Periods; that is, in Periods A and C (with the ultrasound), the mean number of trials was 102 (SD 29), and the average number of practice attempts per session during Periods B and D (without the ultrasound) was 107 (SD 21). The number of practice trials did not differ between Phase 1 (mean 235 practice trials per session) and Phase 2 (mean 184 trials per session, $p = 0.16$), and there was also no difference in number of trials per session between the Prosodic Cueing condition (mean 200 trials) vs. No Prosodic Cueing conditions (mean 219 trials, $p = 0.89$).

Treatment fidelity and reliability

The duration of the 13 min Periods was controlled by use of a timer, thus adherence to this structure was always controlled. In occasional instances in which the family arrived late or the probes took longer than scheduled, the final 13 min block (Period D, with no ultrasound) was cut short.

The second author or a research assistant reviewed recordings of two sessions per participant (once in each condition) to document adherence to the protocol with respect to verbal feedback. In these 16 sessions that were reviewed (totaling 3639 trials), the specified feedback (KP+KR, KR only, or no feedback) was provided on 89% of trials. Deviations from the protocol typically involved providing positive KR ("good," or "mm-hmm") when no feedback was required, or providing only KR when KP+KR was required.

All probes were scored by two listeners (the treating clinician and a second listener who was unaware of which sound patterns were being treated at the time) and averaged. On average across all probes, the treating clinician's scores were 3.6% higher than the second listener. Point-by-point agreement was 86.6% for the first two listeners for probes on which the effect sizes were computed. For probes in which the two listeners disagreed by 20% or more, a third listener also scored the probe from the recording and the final data point therefore reflects the average of three listeners (for 27% of probes). These disagreements typically occurred as participants were progressing in speech accuracy but their productions were inconsistently or "marginally correct".

Treatment side-effects

Although diagnostic ultrasound safety has been documented (AIUM, 2012) and the low-intensity ultrasound used here provides output substantially below levels that could generate bioeffects, we sought to verify that no undesirable side-effects were observed. Both the participant and the parent were independently queried about side-effects. Participants were asked two questions: "When you are using the ultrasound, is there anything that you do not like about it?" and "When you are not here using the ultrasound, have you had any problems with your mouth that you think might be due to the use of the ultrasound?" Parents were asked two questions: "Since the beginning of the study, have you observed any unwanted side-effects?" and "Since the beginning of the study, has your child complained about having trouble eating or pain or discomfort in or around the mouth?"

Data analysis

As this study involves single case design methodology, interpretations of results were based on visual screening of the data as well as effect size calculations. Because there are no single effect size measures that are fully agreed upon in communication disorders research (and minimal guidance from prior research on RSSEs), we include multiple measures for future comparison. The raw percentage increase in accuracy is a value typically used by clinicians. Additionally, a variation on Cohen's d (Cohen, 1998) was selected here: 'd_2,' is relatively conservative measure that has been designed for use in single case research and is calculable when there is zero variance pre-treatment, as occurs when speakers say none of the target sounds correctly in baseline (Beeson & Robey, 2006).. This d_2 measure is calculated as the

difference between post-treatment mean and pre-treatment mean divided by the pooled variance; this measure is calculated using an estimate from only from the data points immediately before and immediately after treatment (i.e., baseline phases) to avoid any possible contamination from treatment during the other treatment phase. Our application of the d_2 measure is conservative in the sense that if one were to include a larger number of zero scores from baseline, the pooled variance would decrease and thus d2 would increase. By including only a small number of those zero-scores in the pooled variance estimate, we do not artificially deflate the pooled variance. In this study, we collected either three or four baseline measures from participants and therefore used the same number of data points from mid and post data collection points. Additionally, percent of non-overlapping data (PND, the percent of post-treatment data points that are higher than the highest pre-treatment data point) was calculated from immediately pre to immediately post each phase of treatment.

Results

Individual participant data are presented in Figure 1 and in Table 3. Figure 1 shows participants' performance from the 25-word probes associated with treated speech sounds across two conditions. In panels A and B of Figure 1 the four participants who received the prosodic condition first followed by the non-prosodic condition are shown. In panels C and D the participants who received the non-prosodic condition first and the prosodic condition second are shown. Boxes on each panel show the period in which treatment was provided on the target with the nominated condition. Panel E demonstrates additional experimental control across four participants (#74, #83, #89, #91) via mean accuracy of sound patterns that were untreated and phonetically unrelated to the treated sound patterns. It should be noted that it was not possible to obtain control data for the other participants as they had no phonemes in error which were unrelated to the two treatment targets.

The primary hypothesis was that children with RSSE's persisting speech sound errors would show improved accuracy with the ultrasound visual feedback. As can be seen in visual inspection of Figure 1, most participants demonstrated a treatment effect (as evidenced by improvements over baseline, coincident with onset of treatment) in one or both phases of ultrasound intervention. Panels A and B present the first four participants. In Panel A, it is clear that participants #86 and #89 show clear acquisition and maintenance of the first treated behavior, whereas #83 and #85 show a more protracted period of improvement that continues after treatment is withdrawn. Panel B shows positive treatment effects for participants #85 and #86; participant #83 shows minimal improvement during probes administered during and immediately after treatment, but she does show evidence of improvement at the two month follow-up. Panel C shows treatment effects for the first treated sound targets for #87 and #91, and Panel D shows treatment effects for the second targets for #74 and #87. Additionally, all but participant #88 showed evidence of longer-term learning: there is retention of treatment effect for 7/8 participants for at least one condition. Some participants showed further evidence of motor learning through continued improvement on treated sound contexts during the two-month follow up (see #83, #85 in panel A, and #87, #91 in panel C). At the two-month follow-up, 11 of 16 treated sound contexts were above 60% accurate. Thus, the first hypothesis was confirmed.

To further support the results from Figure 1, Table 3 presents the raw percent change on mean probe data pre- to post-treatment and two effect size measures for each participant (PND, and d_2). All participants showed change in pre- to post-treatment raw means in at least one phase. PND interpretation using the guidelines provided by Scruggs and Mastropieri (1998) suggests that seven sessions of treatment resulted in 11/16 phases providing effective therapy (>70%), 3/16 phases providing questionable treatment effect (50%–70%) and 2/16 phases providing no treatment effect.

If we arbitrarily define a treatment effect as a raw increase of 20% or more on the probes from immediately pre-treatment to immediately post-treatment, then treatment effects were observed in 12/16 phases (with negligible improvement for both phases for #88, the first phase for #87, and the second phase for #89); however, continued growth was observed for one of these (#87 first phase). Of those phases that showed a treatment effect, d_2 values were all above 1.5.

The second hypothesis was that the addition of prosodic cueing would show increased gains over a non-prosodic cueing condition. Table 4 shows pooled treatment effects for the two conditions and the two phases of treatment. Although there was variable response across the two conditions for the individual participants, combining data across phases there appears to be no difference in outcome between the prosodic and non-prosodic conditions. The 95% confidence intervals overlap for all of the reported effect sizes between the two conditions. Therefore, the second hypothesis was not confirmed. It is evident from the effect sizes and associated confidence intervals that both prosodic and no prosodic cueing conditions resulted in improved speech sound accuracy.

We explored the possibility of an order effect by comparing the effect sizes for phase 1 and phase 2 without regard to treatment condition. There is overlap in the 95% confidence between phase 1 and phase 2 intervals for both the d_2 and the PND values. Therefore, there did not appear to be strong evidence of an order effect.

Experimental control and generalization

For all participants, one or more untreated sound contexts were probed. For four of the participants, an untreated sound context was probed that was phonetically dissimilar to the treated sound context (e.g., word-initial /tʃ/ and post-vocalic /ɹ/ were treated but word-final /θ/ was probed; see Panel E of Fig. 1 for mean data). As can be seen in Figure 1, there was no consistent or appreciable change on the accuracy of these items, providing further support for the interpretation that change observed in treated sound targets were due to the intervention and not generalized improvement (e.g., maturation).

For the other four participants, their speech sound errors were restricted to a narrow class of errors and therefore sound contexts that were monitored but untreated were phonetically similar to treated sound contexts (e.g., word-initial /ɹ/ singleton was treated but word-initial /ɹ/ clusters were probed). We explored whether there was evidence of generalization from treated to phonetically similar untreated sound contexts, as this would enhance the claim that motor *learning* was occurring. As seen in Figure 2, there was clear evidence of generalization from treated to untreated but phonetically related sound contexts in the onset

position. That is, in three participants, treating onset singletons /ɹ, s/ generalized to onset clusters /Cɹ, sC/, and in one participant treating onset clusters /Cɹ/ generalized to onset singletons /ɹ/. Within these four participants, the correlations between the treated and the untreated onset singleton-cluster pairs shown in Figure 2 were high, ranging from $r = 0.85 – 0.98$, indicating close correspondence between treated and untreated but similar targets in the onset position of words.

The /s/ and /ɹ/ tokens in the 15 sentence imitation task were scored from audio recordings by two research assistants. The average of the two listeners was computed for each time point. A Wilcoxon Signed-Rank test was used to compare pre- and post-treatment scores. Accuracy on the sentence task was significantly higher for post-treatment (mean 57% SD 20%, median 61%) than for pre-treatment (mean 42% SD 10%, median 37%, Z= 1.96; $p = 0.025$ one-tailed).

Treatment side-effects

In response to the queries about side-effects, all parents reported that there were no observed side effects and that their child had not complained about the ultrasound. When the participants were asked if there were problems with their mouths that might be due to the use of the ultrasound, all 8 participants said "no." When asked if there was anything they did not like about the treatment, the following was reported: the ultrasound gel was "gooey" or "cold," or the probe was "sometimes annoying" when putting pressure under the chin.

Discussion

This Phase II study used single subject experimental methods to study the effects of ultrasound visual feedback for individuals with RSSEs. Because these participants had persisting speech errors, they were unlikely to resolve their speech errors without treatment, and this was confirmed by minimal change in pre-treatment baselines and by the stability of untreated sounds (see Figure 1, Panel E). Each participant received 14 treatment sessions, with seven sessions addressing each of two target sound contexts. The primary hypothesis that a motor learning-based ultrasound visual feedback treatment would result in improved speech sound accuracy was confirmed through replication across multiple participants and multiple treatment targets. Most participants showed evidence of a treatment effect (increase over baseline) with at least one of their treated sound patterns, although not every participant responded equally. The results of the study generally support evidence from previous Phase I and Phase II research demonstrating positive outcomes with ultrasound feedback for individuals with RSSEs (Adler-Bock et al., 2007; Modha et al., 2008; Shawker & Sonies, 1985), CAS (Preston et al., 2013), acquired apraxia of speech (Preston & Leaman, 2014) and hearing impairment (Bacsfalvi, 2010; Bacsfalvi, Bernhardt, & Gick, 2007; Bernhardt, Gick, Bacsfalvi, & Ashdown, 2003). The present study extends prior research by including greater experimental control (e.g., multiple baseline with replication, untreated control items) and by incorporating a structured, replicable protocol based on principles of motor learning. The observed treatment outcomes support the exploration of future Phase III experimental studies.

The secondary hypothesis – that practice variability through prosodic cueing would enhance motor learning – was not supported by the data. All possible combinations of outcomes were observed: some participants showed equally strong long-term responses to both prosodic and non-prosodic cueing (e.g., participants #86, #87); others showed minimal response to both conditions (e.g., #88); some had clearer benefit from prosodic cueing over no prosodic cueing (e.g., #89); and some participants showed greater learning with no prosodic cueing (e.g., #91). Thus, as implemented here, the prosodic cueing did not appear to provide a robust addition to the ultrasound treatment for these participants. It is possible that this type of prosodic cueing was not sufficiently robust to facilitate additional learning, as it involved practicing target words and phrases three different ways (statement, command, question). It is noteworthy, however, that the current treatment program includes other elements of practice variability built in to the protocol, such as practicing two variants of the sound context (e.g., treating /ɹe/ and /ɹo/ to address initial /ɹ-/) and practicing multiple words within a session. These other elements may have allowed for sufficient practice variability and the prosodic cues may have been superfluous. Other individuals, such as those with CAS, might benefit from explicit attempts to pair articulatory and prosodic cues (cf. Preston et al., 2013 in which prosodic cueing was included in a less explicit manner).

Ultrasound visual feedback provides the participant with explicit KP. This information allows the clinician to cue elements of articulatory movement (e.g., tongue root retraction, tongue dorsum lowering, tongue grooving) that cannot be as easily described to the participant without ultrasound. These elements are believed to enhance learning because they provide a more explicit reference for both the client and clinician when teaching a new GMP. However, it is important to note that the treatment structure was built around several principles of motor learning (cf. Maas et al., 2008; Schmidt & Lee, 2011), not just the use of ultrasound. For example, many practice trials were elicited yielding a relatively high dose, which research has shown to be helpful in teaching new speech sounds (Edeal & Gildersleeve-Neumann, 2011). Additionally, as the participants achieved greater success, the complexity of the stimuli was designed to increase (i.e., from syllables to monosyllabic words, multisyllabic words, and phrases), the clinician's feedback frequency reduced, and the feedback type changed from primarily KP+KR to primarily KR. Thus, the structure of each session was designed to work up to, but not beyond, the participant's challenge point (Guadagnoli & Lee, 2004; Rvachew & Brosseau-Lapré, 2012). It should therefore be noted that this theoretically-motivated *combination* of principles may contribute to improvement, and the present study was not designed to isolate a single factor responsible for treatment gains. However, clinicians can be moderately confident that a short-term treatment program, as implemented here, can facilitate improvements in speech sound accuracy for some individuals with RSSEs. Future studies that compare the current set of treatment principles both with and without ultrasound visual feedback would be of value. Additionally, modifications to the current treatment protocol could be made to further align with principles of motor learning, such as transitioning from blocked to random practice schedules, transitioning from immediate to delayed feedback, or including more perceptual training and self-evaluation (Murray, McCabe, & Ballard, 2012).

Although current clinical practices vary, some individuals with RSSEs may be dismissed from treatment because of a failure to respond to previous therapies. All but one of these participants (participant #85) had previous unsuccessful therapy addressing the target sounds. Although prior treatment approaches for these participants were likely eclectic and cannot be evaluated, the study provides evidence that individuals with RSSEs may have sufficient neuro-phonetic plasticity to be able to modify their speech sound production. The participants in the present study had no other concomitant speech or language impairments (i.e., no current language or motor speech disorders), and all of the 10–13 year old children showed a response to treatment for at least one treated speech sound; this profile may therefore reflect "good" candidates for treatment. The rate or magnitude of response observed here might vary for other clinical populations (e.g., those with hearing impairment, dysarthria, CAS, language impairment, etc.), but the study provides clear experimental evidence that some individuals with persisting speech sound errors can respond to treatment. Thus, ultrasound visual feedback should be considered one viable treatment option for individuals who may not have responded well to other approaches.

Participant #88 showed the poorest response to treatment overall and he was several years older than the other participants (i.e., college-age). Although some improvements were observed during treatment, motor learning was not evident for this participant. Subjectively, there did not appear to be substantial differences in this participant's level of motivation or participation in the treatment program, and it might therefore be that individuals with such longstanding speech errors would require more treatment to facilitate generalization and retention. Other participants (e.g., participant #91's /-ɹ/) improved during the treatment sessions but did not generalize to the probes. Future studies could explore modifications to the current treatment program, such as the use of ultrasound visual feedback in greater cumulative dose (i.e., more than 14 total sessions, more successful practice attempts) and with greater intensity (i.e., more than two sessions per week) for such individuals.

Despite the individual variation in the magnitude of response to the treatment, the raw percent change and the overall effect sizes are encouraging. Most participants showed significant gains with just seven treatment sessions addressing a single target. Generalization effects in onset position (from trained singletons to clusters, and from trained clusters to singletons) were relatively robust, strengthening the clinical outcomes of the study. Moreover, the improvement in sentence-level accuracy also suggests generalization effects beyond the word level. The two-month follow-up data also support the notion that retention (and in some cases, continued improvement) was evident. Thus, generalization and retention strengthen the claim that the treatment program facilitated motor learning. Following the 14 sessions, two participants (#86, #87) returned to their school-based clinicians and were dismissed from treatment, and at the two-month follow-up participant #83 reported that she was scheduled to be dismissed from treatment in the near future. Additionally, #85 had been deemed ineligible to receive speech services through his school (due to strong academic skills), but at the two-month follow-up his mother felt that there was no need to further pursue treatment.

Finally, although the primary goal of the study was to evaluate treatment effects, we asked participants and their parents about any unwanted side-effects. Although this type of

diagnostic ultrasound use has been extensively studied with no known adverse effects (AIUM, 2012; Barnett et al., 2000; Epstein, 2005), prior studies using ultrasound in speech therapy have not directly reported side-effects. In this study, there were no reported adverse effects, and the only comments provided by the participants were related to the ultrasound gel ("cold" or "gooey") or the discomfort of placing and maintaining the ultrasound probe beneath the chin during therapy. Aside from these minor nuisances, there does not appear to be any obvious adverse effect of ultrasound in speech therapy.

Caveats and Limitations

The treatment effects here are likely the results of the collection of principles applied, not just the implementation of KP with the ultrasound. That is, several principles of motor learning are included in the treatment protocol, such as a high trial rate, increasing levels of complexity, fading feedback, etc. Modifications of the protocol could result in changes in the outcomes. For example, fewer practice trials or practice only on monosyllabic items might not result in the same level of improvement.

There was sufficient experimental control in this study to demonstrate that the treatment program resulted in greater improvement over no treatment. However, no direct comparison to other treatment methods can be made. Future studies comparing a motor-based ultrasound visual feedback treatment to other treatment approaches in Phase III studies would be of clinical value.

Summary and Conclusions

The study provides evidence that ultrasound visual feedback in the context of a motor learning program can be used to facilitate more accurate speech sound production in individuals with RSSEs. Beyond the effects of ultrasound visual feedback, prosodic cueing was not observed to have a robust impact on motor learning in this study. In sum, for some individuals whose speech sound errors have persisted for several years, a short period of ultrasound visual feedback treatment in the context of a motor learning program may be an effective approach to remediation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

Adams SG, Page AD. Effects of selected practice and feedback variables on speech motor learning. Journal of Medical Speech-Language Pathology. 2000; 8(4):215–220.

Adler-Bock M, Bernhardt B, Gick B, Bacsfalvi P. The use of ultrasound in remediation of north american english /r/ in 2 adolescents. American Journal of Speech-Language Pathology. 2007; 16(2):128–139.10.1044/1058-0360(2007/017) [PubMed: 17456891]

AIUM. Prudent use and clinical safety. 2012. Retrieved Nov. 3, 2013, from http://www.aium.org/officialStatements/34

American Speech-Language-Hearing Association. 2010 schools survey report: SLP caseload characteristics. 2010. Available from www.asha.org/research/memberdata/SchoolsSurvey.htm

Bacsfalvi P. Attaining the lingual components of /r/ with ultrasound for three adolescents with cochlear implants. Journal of Speech-Language Pathology and Audiology. 2010; 34(3):206–217.

Bacsfalvi P, Bernhardt BM. Long-term outcomes of speech therapy for seven adolescents with visual feedback technologies: Ultrasound and electropalatography. Clinical Linguistics & Phonetics. 2011; 25(11–12):1034–1043.10.3109/02699206.2011.618236 [PubMed: 22106893]

Bacsfalvi P, Bernhardt BM, Gick B. Electropalatography and ultrasound in vowel remediation for adolescents with hearing impairment. International Journal of Speech-Language Pathology. 2007; 9(1):36–45.10.1080/14417040601101037

Ballard KJ, Maas E, Robin DA. Treating control of voicing in apraxia of speech with variable practice. Aphasiology. 2007; 21(12):1195–1217.

Barnett SB, Ter Haar GR, Ziskin MC, Rott HD, Duck FA, Maeda K. International recommendations and guidelines for the safe use of diagnostic ultrasound in medicine. Ultrasound in Medicine & Biology. 2000; 26(3):355–366. [PubMed: 10773365]

Beeson PM, Robey RR. Evaluating single-subject treatment research: Lessons learned from the aphasia literature. Neuropsychology Review. 2006; 16(4):161–169. [PubMed: 17151940]

Bernhardt B, Gick B, Bacsfalvi P, Ashdown J. Speech habilitation of hard of hearing adolescents using electropalatography and ultrasound as evaluated by trained listeners. Clinical Linguistics & Phonetics. 2003; 17(3):199–216.10.1080/0269920031000071451 [PubMed: 12858839]

Campbell F, Gick B, Wilson I, Vatikiotis-Bateson E. Spatial and temporal properties of gestures in north american english /r/. Language & Speech. 2010; 53(Pt 1):49–69. [PubMed: 20415002]

Cohen, J. Statistical power analysis for the behavioral sciencies. 2. Hillsdale, NJ: Erlbaum; 1998.

Dagenais PA. Electropalatography in the treatment of articulation/phonological disorders. Journal of Communication Disorders. 1995; 28(4):303–329. 0021992495000591 [pii]. [PubMed: 8576412]

Dent H, Gibbon F, Hardcastle WJ. The application of electropalatography (EPG) to the remediation of speech disorders in school-aged children and young adults. International Journal of Language & Communication Disorders. 1995; 30:264–277.

Dodd, B.; Hua, Z.; Crosbie, S.; Holm, A.; Ozanne, A. Diagnostic Evaluation of Articulation and Phonology (DEAP). London: Psych-Corp; 2002.

Dunn, LM.; Dunn, DM. Peabody Picture Vocabulary Test. 4. Minneapolis, MN: Pearson; 2007.

Edeal DM, Gildersleeve-Neumann CE. The importance of production frequency in therapy for childhood apraxia of speech. American Journal of Speech-Language Pathology. 2011; 20(2):95–110.10.1044/1058-0360(2011/09-0005) [PubMed: 21330650]

Epstein MA. Ultrasound and the IB. Clinical Linguistics & Phonetics. 2005; 19(6–7):567–572.10.1080/02699200500113889 [PubMed: 16206484]

Fey, ME.; Finestack, LH. Research and development in child language intervention: A 5-phase model. In: Schwartz, RG., editor. Handbook of child language disorders. Vol. 513–529. New York, NY: Psychology Press; 2009.

Flipsen P. Articulation rate and speech-sound normalization failure. Journal of Speech, Language, and Hearing Research. 2003; 46:724–737.

Gibbon F, Hardcastle W. Articulatory description and treatment of "lateral /s/" using electropalatography: A case study. British Journal of Disorders of Communication. 1987; 22:203–217. [PubMed: 3504703]

Goldman, R.; Fristoe, M. Goldman Fristoe Test of Articulation. 2. Circle Pines, MN: AGS; 2000.

Gruber FA. Probability estimates and paths to consonant normalization in children with speech delay. Journal of Speech, Language, and Hearing Research. 1999; 42:448–459.

Guadagnoli MA, Lee TD. Challenge point: A framework for conceptualizing the effects of various practice conditions in motor learning. Journal of Motor Behavior. 2004; 36(2):212–224. [PubMed: 15130871]

Hall KG, Magill RA. Variability of practice and contextual interference in motor skill learning. Journal of Motor Behavior. 1995; 27(4):299–309.10.1080/00222895.1995.9941719 [PubMed: 12529226]

Irwin JV, Huskey R, Knight N, Oltman S. A longitudinal study of the spontaneous remission of articulatory defects of 1665 school children in grades 1, 2, and 3: Ii. The sample. Acta Symbolica. 1974; 5(2):1–7.

Lipetz HM, Bernhardt BM. A multi-modal approach to intervention for one adolescent's frontal lisp. Clinical Linguistics & Phonetics. 2013; 27(1):1–17.10.3109/02699206.2012.734366 [PubMed: 23237414]

Maas E, Robin DA, Austermann Hula SN, Freedman SE, Wulf G, Ballard KJ, Schmidt RA. Principles of motor learning in treatment of motor speech disorders. Am J Speech Lang Pathol. 2008; 17(3): 277–298.10.1044/1058-0360(2008/025) [PubMed: 18663111]

McAllister Byun T, Hitchcock ER. Investigating the use of traditional and spectral biofeedback approaches to intervention for /r/ misarticulation. American Journal of Speech-Language Pathology. 2012; 21(3):207–221.10.1044/1058-0360(2012/11-0083) [PubMed: 22442281]

Modha G, Bernhardt B, Church R, Bacsfalvi P. Ultrasound in treatment of /r/: A case study. International Journal of Language & Communication Disorders. 2008; 43(3):323–329. [PubMed: 17852527]

Murray E, McCabe P, Ballard KJ. A comparison of two treatments for childhood apraxia of speech: Methods and treatment protocol for a parallel group randomised control trial. BMC Pediatrics. 2012; 12:112.10.1186/1471-2431-12-112 [PubMed: 22863021]

Preston JL, Brick N, Landi N. Ultrasound biofeedback treatment for persisting childhood apraxia of speech. American Journal of Speech-Language Pathology. 2013; 22(4):627–643.10.1044/1058-0360(2013/12-0139) [PubMed: 23813207]

Preston JL, Edwards ML. Phonological processing skills of adolescents with residual speech sound errors. Language, Speech and Hearing Services in Schools. 2007; 38:297–308.

Preston JL, Leaman M. Ultrasound visual feedback for acquired apraxia of speech: A case report. Aphasiology. 2014; 28(3):278–295.

Robey RR. A five-phase model for clinical-outcome research. Journal of Communication Disorders. 2004; 37(5):401–411. [PubMed: 15231420]

Rvachew, S.; Brosseau-Lapré, F. Developmental phonological disorders: Foundations of clinical practice. San Diego: Plural Publishing; 2012.

Rvachew S, Hodge M, Ohberg A. Obtaining and interpreting maximum performance tasks from children: A tutorial. Journal of Speech-Language Pathology and Audiology. 2005; 29(4):146–157.

Sax MR. A longitudinal study of articulation change. Language, Speech & Hearing Services in Schools. 1972; 3(1):41–48.

Schmidt, RA.; Lee, TD. Motor control and learning: A behavioral emphasis. 2. Human Kinetics; 2011.

Scruggs TE, Mastropieri MA. Summarizing single-subject research: Issues and applications. Behavior Modification. 1998; 22(3):221–242.10.1177/01454455980223001 [PubMed: 9722473]

Secord, WA.; Boyce, SE.; Donohue, JS.; Fox, RA.; Shine, RE. Eliciting sounds: Techniques and strategies for clinicians. 2. Clifton Park, NY: Thomson Delmar Learning; 2007.

Semel, E.; Wiig, EH.; Secord, WA. Clinical evaluation of language fundamentals. 4. San Antonio, TX: Harcourt Assessment; 2003.

Shawker TH, Sonies BC. Ultrasound biofeedback for speech training: Instrumentation and preliminary results. Investigative Radiology. 1985; 20(1):90–93. [PubMed: 3980183]

Shriberg LD. A response evocation program for /er/. Journal of Speech & Hearing Disorders. 1975; 40(1):92–105. [PubMed: 1123932]

Shriberg LD. An intervention procedure for children with persistent /r/ errors. Language, Speech and Hearing Services in Schools. 1980; 11(2):102–110.

Shriberg LD, Austin D, Lewis BA, McSweeny JL, Wilson DL. The speech disorders classification system (sdcs): Extensions and lifespan reference data. Journal of Speech, Language & Hearing Research. 1997; 40(4):723–740.

Shriberg LD, Fourakis M, Hall SD, Karlsson HB, Lohmeier HL, McSweeny JL, Wilson DL. Extensions to the speech disorders classification system (SDCS). Clinical Linguistics & Phonetics. 2010; 24(10):795–824.10.3109/02699206.2010.503006 [PubMed: 20831378]

Shriberg LD, Kwiatkowski J. Phonological disorders ii: A conceptual framework for management. Journal of Speech and Hearing Disorders. 1982; 47:242–256. [PubMed: 7186560]

Shuster LI, Ruscello DM, Smith KD. Evoking [r] using visual feedback. American Journal of Speech-Language Pathology. 1992; 1(3):29–34.

Strand EA, Debertine P. The efficacy of integral stimulation intervention with developmental apraxia of speech. Journal of Medical Speech Language Pathology. 2000; 8(4):295–300.

Strand EA, Stoeckel R, Baas B. Treatment of severe childhood apraxia of speech: A treatment efficacy study. Journal of Medical Speech Language Pathology. 2006; 14(4):297.

Van Riper, C.; Erickson, RL. Speech correction: An introduction to speech pathology and audiology. Allyn and Bacon; 1996.

Wagner, RK.; Torgesen, JK.; Rashotte, CA.; Pearson, NR. Comprehensive test of phonological processing. 2. Austin, TX: Pro-Ed; 2013.

Wechsler, D. Wechsler abbreviated scales of intelligence. 2. Pearson; 2011.

Williams AL. Generalization patterns associated with training least phonological knowledge. Journal of Speech & Hearing Research. 1991; 34(4):722–733. [PubMed: 1956180]

Williams, KT. Expressive Vocabulary Test. 2. Pearson; 2007.

Wong AYH, Ma EPM, Yiu EML. Effects of practice variability on learning of relaxed phonation in vocally hyperfunctional speakers. Journal of Voice. 2011; 25(3):e103–e113. [PubMed: 20456910]

Wulf G, Schmidt RA. Variability of practice and implicit motor learning. Journal of Experimental Psychology: Learning, Memory, and Cognition. 1997; 23(4):987–1006.10.1037/0278-7393.23.4.98
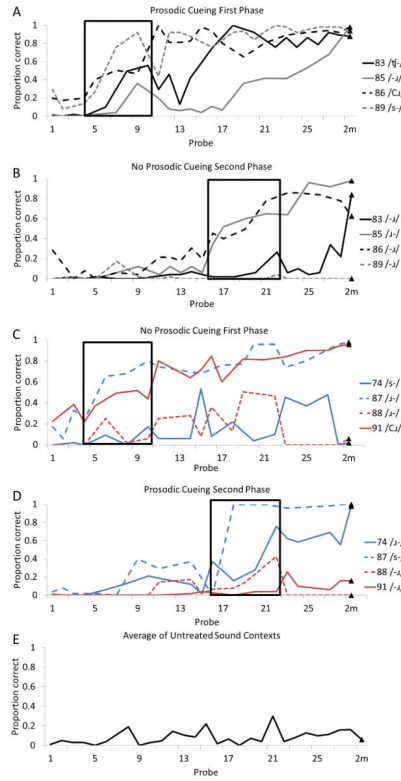
Figure 1.

Probe data for eight participants

Notes. Panels A and B show four participants who were treated with Prosodic Cueing first (Panel A) followed by the No Prosody condition (Panel B) on a different sound context. Panels C and D represent four different participants who were treated with the No Prosody condition first (Panel C) followed by the Prosodic Cueing condition (Panel D). Sound patterns that were phonetically dissimilar from the treated sound contexts were probed for four participants and the average of these are shown in Panel E. Boxes represent sessions in which treatment occurred. Triangles represent accuracy at a two-month follow-up session.

Figure 2.
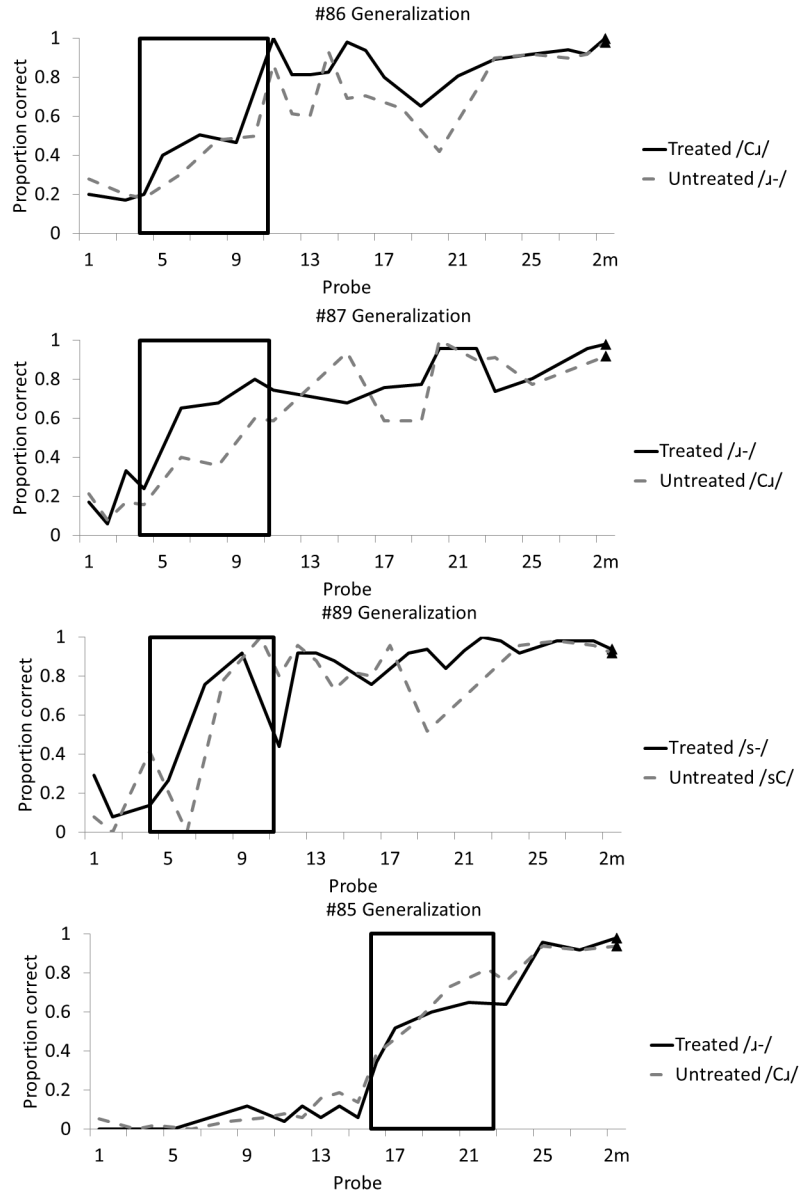Generalization effects among onset singletons and onset clusters
Notes. Each of the four panels represents probe data from a different participant. Solid black lines represent the treated onset (singleton or cluster) and dashed gray lines represent the untreated onset (singleton or cluster). Boxes represent sessions in which treatment occurred. Triangles represent accuracy at a two-month follow-up session.

Table 1

Participant characteristics

Participant	Age	Gender	WASI-II PRI	GFTA-2 SS	CTOPP-2 PA SS	CELF-4 Recalling Sent SS	CELF-4 Formulated Sent SS	PPVT-4 SS	EVT-2 SS
74	10	F	39	40	88	7	12	95	92
83	12	F	9	<40	92	11	12	119	115
85	12	M	19	71	90	10	14	117	120
86	10	M	73	66	94	8	11	98	97
87	10	M	61	53	118	9	12	112	110
88	20	M	88	<40	N/A	11	12	126	114
89	13	M	42	<40	112	11	13	107	106
91	12	M	77	71	100	9	9	100	108

Notes. WASI PRI= Wechsler Abbreviated Scales of Intelligence Perceptual Reasoning Index; GFTA-2 = Goldman-Fristoe Test of Articulation-2; SS= Standard Score; CTOPP-2 PA= Comprehensive Test of Phonological Processing-2 Phonological Awareness Composite; CELF-4 = Clinical Evaluation of Language Fundamentals; PPVT-4= Peabody Picture Vocabulary Test-4; EVT-2= Expressive Vocabulary Test-2. N/A = not administered.

Table 2

Structured Practice and Feedback Schedule (with six trials per block)

Level	Trials with KP+KP	Trials with KR only	Trials with no Feedback	Example
Syllables	5	0	1	/re/
Monosyllabic words	3	2	1	Rain
Multisyllabic words	2	2	2	Rainbow
Phrases	1	2	3	Rainbows in the sky
Cloze Phrases	1	2	3	Rainbows are ___

Notes: Trials at each level were presented in blocks of 6. Step up criteria to the next level was 5 of 6 correct in a block. The same practice schedule was used for Periods A through D.

Table 3

Sound contexts treated under Prosodic Cueing and No Prosodic Cueing conditions with effect size calculated on 3 or 4 probes pre and post treatment in each phase

Participant	Phase 1 Condition	Context (variants)	Raw % change	PND	d_2	Phase 2 Condition	Context (variants)	Raw % change	PND	d_2
74	No Prosody	Initial /s/ (/si, saɪ/)	21	83	1.51	Prosody	Initial /ɹ/ (/ɹe, ɹoɪ/)	44	100	4.28
83	Prosody	Initial /tʃ/ (/tʃi, tʃɛ/)	29	100	2.42	No Prosody	Vocalic /ɹ/ (/aɪ, ɔɪ/)	2	50	0.84
85	Prosody	Vocalic /ɹ/ (/aɪ, ɪɪ/)	29	100	5.77	No Prosody	Initial /ɹ/ (/ɹe, ɹoɪ/)	73	67	3.59
86	Prosody	/ɹ/ clusters (/fɹɪ, θɹɪ/)	82	100	8.61	No Prosody	Vocalic /ɹ/ (/aɪ, ɪɪ/)	55	100	8.04
87	No Prosody	Initial /ɹ/ (/ɹe, ɹoɪ/)	7	100	6.69	Prosody	Initial /s/ (/si, saɪ/)	76	100	7.02
88	No Prosody	Initial /ɹ/ (/ɹe, ɹoɪ/)	20	67	3.41	Prosody	Vocalic /ɹ/ (/aɪ, ɪɪ/)	−9	0	−2.2
89	Prosody	Initial /s/ (/si, saɪ/)	82	100	3.04	No Prosody	Vocalic /ɹ/ (/aɪ, ɪɪ/)	0	0	0
91	No Prosody	/ɹ/ clusters (/fɹɪ, kɹɪ/)	44	100	4.74	Prosody	Vocalic /ɹ/ (/aɪ, ɪɪ/)	13	100	2.36
MEAN (SD)			39 (25)	93 (12)	4.52 (2.39)			32 (34)	64 (44)	2.99 (3.48)

Notes. PND= percent non overlapping data, d_2 = mean difference of the immediate 3–4 pre-treatment values and 3–4 post treatment values, divided by the pooled standard deviation of these pre- and post-treatment values.

Table 4

Treatment effect by condition without reference to order and by order without reference to condition

	Raw % change		PND		d_2	
	mean (SD)	95% CI	mean (SD)	95% CI	mean (SD)	95% CI
Prosodic condition	37 (33)	14–60	88 (35)	63–100	3.9 (3.3)	1.6–6.2
Non prosodic condition	34 (27)	15–52	71 (34)	47–94	3.6 (2.8)	1.7–5.6
Phase 1	39 (25)	22 – 56	93 (12)	85–100	4.5 (2.4)	2.9–6.2
Phase 2	32 (34)	8 – 56	65 (44)	34–95	3.0 (3.5)	0.6–5.4

Notes. PND= percent non overlapping data, d_2 = mean difference of the immediate 3–4 pre-treatment values and 3–4 post treatment values, divided by the pooled standard deviation of these pre- and post-treatment values.