AJSLP

Clinical Focus

Intensive Treatment for Persisting Rhotic Distortions: A Case Series

Jonathan L. Preston[a,b] and Megan C. Leece[a]

Purpose: The study explored changes in accuracy of American English rhotics as a result of an intensive 1-week therapy program for adolescents and young adults with residual speech sound errors that had not resolved with previous therapy.
Method: Four case reports are presented of individuals aged 13, 17, 21, and 22 years with residual /ɹ/ distortions. Each participant attended a 1-week intensive program consisting of pretreatment assessments, 14 hr of therapy, and posttreatment assessment. Treatment sessions included structured motor-based practice, ultrasound visual feedback of the tongue, and auditory speech perception training. To assess generalization, untreated words and sentences with rhotics were recorded before and after therapy; these were rated by listeners who were blind to when the recordings were taken.
Results: All participants showed measurable and statistically significant improvement in speech sound accuracy. Averaged across the 4 participants, rhotic accuracy at the word level improved from 35% to 83%. At the sentence level, rhotic accuracy increased from 11% pretreatment to 66% posttreatment in 1 week.
Conclusion: The promise of an intensive treatment program that includes motor-based practice, biofeedback, and auditory perception training is illustrated by the case presentations in which substantial improvements in speech sound accuracy were observed.
Supplemental Materials: https://doi.org/10.23641/asha.5561254

Some individuals with residual speech sound errors (RSSEs) may experience limited response to traditional speech therapy methods, resulting in errors that persist into adolescence and adulthood. RSSEs, which are speech sound errors that persist beyond approximately 8–9 years old, are present in approximately 1%–2% of the population (Flipsen, 2015). RSSEs may be associated with histories of speech sound disorders (e.g., phonological disorders or childhood apraxia of speech) or may be the result of typical speech sound acquisition with the exception of distortions of later developing sounds (Flipsen, 2015). Not all individuals with RSSEs are necessarily eligible for speech therapy services through local educational agencies, particularly when they fail to show a significant academic impact as a result of their RSSEs; some individuals who can access services may receive treatments that fail to yield measurable progress in the accuracy or clarity of the client's speech, despite the best efforts of speech-language pathologists (SLPs). Although academic performance may not always suffer, social consequences of RSSEs in and out of school remain quite real, such as negative peer perceptions, bullying, and limited participation in extracurricular activities (e.g., theater) or limited employment opportunities (Hitchcock, Harel, & McAllister Byun, 2015; McCormack, McLeod, McAllister, & Harrison, 2009; Silverman & Paulus, 1989). Thus, for those individuals with RSSEs who wish to seek alternative treatments when previous clinical efforts have been insufficient, efficacious therapy options must be explored.

Most RSSEs involve distortions of later developing sounds such as /ɹ, l, s, z, ʃ/, although distortions of /ɹ/ and the rhotic vowels /ɝ, ɚ/ appear to be the most common in American English (Shriberg, 2009). It is difficult to characterize existing therapeutic approaches as there are few recent observational studies characterizing current clinical practices for RSSEs in schools or clinics (only laboratory research); anecdotal reports suggest that most approaches to treatment for RSSEs are eclectic and highly individualized. Despite the lack of formal descriptions of current clinical practices, it is plausible that treatment approaches tend to be motor based (rather than phonological). In addition, because individuals with RSSEs have few sounds in error, the severity of any observed impairment may often be viewed as mild; thus, services probably tend to be provided less rather than more frequently, perhaps at a rate

[a]Department of Communication Sciences and Disorders, Syracuse University, NY
[b]Haskins Laboratories, New Haven, CT
Correspondence to Jonathan L. Preston: jopresto@syr.edu

Editor: Krista Wilkinson
Associate Editor: Rebecca McCauley
Received December 1, 2016
Revision received April 25, 2017
Accepted June 13, 2017
https://doi.org/10.1044/2017_AJSLP-16-0232

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

of approximately once per week (Brandel & Frome Loeb, 2011).

When more traditional intervention models have not been successful for clients with RSSEs, alternatives to some of these existing strategies may include (a) implementing practice within a theoretical framework that incorporates several principles of motor learning (PMLs; e.g., massed vs. distributed practice, blocked vs. random practice, rate and type of feedback delivery), (b) using visual feedback of the articulators to teach clients how to alter articulatory movements for target sounds, (c) modifying the intensity of service delivery, and (d) enhancing the client's understanding of auditory goals for speech sounds with sound-specific auditory perception training. Some of these strategies are likely included in traditional speech therapy to various degrees, but this study provides an exploration of a program for RSSEs that explicitly incorporates these four primary elements.

Principles of Motor Learning

Several PMLs have been derived from studies of nonspeech motor learning, offering guidance on conditions of motor practice and feedback that differentially influence acquisition versus learning of a motor skill (Maas et al., 2008; Schmidt & Lee, 2011). Acquisition, which represents motor performance during practice (i.e., during therapy), is distinct from learning, which is the ability to retain and generalize a motor skill. Several PMLs have recently been applied to speech motor learning in individuals with speech sound impairments and may serve as a useful framework for remediation of RSSEs. For individuals with RSSEs, treatment may be structured to transition from an initial focus on acquisition to a later focus on learning.

With respect to feedback during motor practice, acquisition may be enhanced by providing knowledge of performance (KP) feedback (i.e., specific information about the detailed aspects of a movement that is executed). There is evidence from the nonspeech motor learning literature suggesting that KP feedback may facilitate learning movement patterns for which the task goals are unclear to the learner (Newell, Carlton, & Antoniou, 1990). Thus, for individuals for whom acquisition of a speech movement is problematic, KP feedback may be beneficial early in the therapeutic process. However, learning may be facilitated through knowledge of results (KR) feedback (i.e., information about correctness), and it may be appropriate to transition from KP to KR feedback to facilitate retention and generalization (Maas et al., 2008; Schmidt & Lee, 2011). In addition, based on predictions from schema-based motor learning theory, frequent and immediate feedback may be useful early in practice to facilitate acquisition, but a transition to less frequent and delayed feedback may be useful to facilitate learning (Maas et al., 2008).

The theoretical predictions of PMLs also provide guidance on how to practice a motor skill (Hitchcock & McAllister Byun, 2015; Maas et al., 2008). Practicing simple targets may increase performance during treatment, but practicing complex targets may be most useful for motor learning. With respect to speech, a progression from simple targets (such as syllables or words) to complex targets (such as multisyllabic words, phrases, and sentences) may be necessary to support clients in first acquiring and then generalizing/retaining target movements, and this is inherent in many speech therapy programs. In addition, practice in blocks (i.e., repeating the same movement multiple times in a row) may aid acquisition, whereas random practice is hypothesized to lead to motor learning. Furthermore, constant practice (practicing the movement the same way) may aid acquisition, whereas variable practice (practicing the movement different ways) may facilitate motor learning (Skelton & Hagopian, 2014). When structuring treatment, these principles may provide guidance on how to transition from an acquisition-focused approach to a learning-focused approach.

Some of these principles are inherent in traditional speech therapy (e.g., increasing stimulus complexity), but other principles may be included less systematically. This study utilizes a structured treatment involving chaining of syllables, an approach that has been adapted from previous research (Preston et al., 2014; Preston, Leece, & Maas, 2016, 2017); many of the practice and feedback parameters are constructed with the aforementioned principles in mind. Several aspects of the treatment are designed to facilitate acquisition of the speech sound in error, and there are also aspects of the training that involve a progression toward PML (e.g., changing feedback from primarily KP to KR, reducing the frequency of feedback, increasing target complexity, and randomizing practice). As opposed to some traditional approaches to speech therapy that implement adaptations only at the start of a new session (e.g., transitioning from practicing words to phrases), the present PML approach incorporates an adaptive framework in which elements of practice and feedback are constantly modified based on the client's performance to work at a level that will maximize learning (cf. Guadagnoli & Lee, 2004; Hitchcock & McAllister Byun, 2015).

Visual Feedback

Another modification to traditional therapeutic programs includes the addition of visual feedback, which involves instrumentation to provide real-time information about speech. The underlying assumption is that visual information about speech may make transient articulatory movements more overt for both the client and the clinician (Huang, Wolf, & He, 2006). Visual feedback of the articulators can be an added form of KP feedback (in addition to the verbal KP feedback provided by the clinician) and therefore may aid in acquisition of target motor behaviors (Newell et al., 1990). Several visual feedback options are available for treating RSSEs that affect rhotic production, including acoustic feedback, which can be used to present real-time spectrographic or spectral information related to formant patterns for /ɹ/ (e.g., McAllister Byun & Hitchcock, 2012; McAllister Byun & Campbell, 2016; Ruscello, 1995; Shuster, Ruscello, & Smith, 1992); electropalatography, which can show contact of the lateral margins of the tongue against the molars for /ɹ/ (e.g., Fabus et al., 2015; Gibbon & Lee, 2015; Schmidt, 2007); and ultrasound, which shows real-time images of tongue shapes (e.g., Adler-Bock, Bernhardt, Gick, & Bacsfalvi, 2007; Preston et al., 2014; Preston, Leece, & Maas, 2016, 2017). Ultrasound, which is used in this study, can be used to cue changes in tongue shape or position for a variety of lingual speech sounds (e.g., Bacsfalvi & Bernhardt, 2011; Cleland, Scobbie, & Wrench, 2015), and it is particularly well suited to teach lingual configurations for /ɹ/. For example, a correct /ɹ/ often involves an elevated tongue tip/blade, a low dorsum, tongue root retraction toward the pharynx, and elevation of the lateral margins of the tongue with a midline groove; a distorted /ɹ/ may involve a low tongue blade, a high tongue dorsum, a lack of tongue root retraction, and no midline grooving (Bacsfalvi, 2010; Boyce, 2015; Klein, McAllister Byun, Davidson, & Grigos, 2013; Preston et al., 2017). During treatment, ultrasound can be used to point out these aspects of lingual articulation to enhance clients' understanding of the intended and actual tongue movements. Visualization of the articulators enables cues to be more specific and concrete (e.g., "try to raise this part of your tongue a little higher") and may aid clients' understanding of accurate articulatory requirements of speech sounds.

The present investigation focuses on ultrasound visual feedback because a number of case reports and small-scale experimental studies have shown that it may successfully facilitate improvements in /ɹ/ accuracy for individuals with RSSEs (Adler-Bock et al., 2007; Bernhardt et al., 2008; McAllister Byun, Hitchcock, & Swartz, 2014; Modha, Bernhardt, Church, & Bacsfalvi, 2008; Preston et al., 2014, 2017). For example, McAllister Byun et al. (2014) reported a mean improvement of consonantal and vocalic /ɹ/ of 54% in 17 sessions of ultrasound biofeedback therapy for four participants aged 9–15 years. Adler-Bock et al. (2007) described two participants, aged 12 and 14 years, whose /ɹ/ improved after 14 treatment sessions with ultrasound biofeedback (improvements of 48% and 62% at the phrase level, respectively). Modha et al. (2008) also reported significant improvement in listeners' ratings of /ɹ/ accuracy, along with corresponding acoustic changes, in a 13-year-old boy. Furthermore, in two studies reporting on 20 participants aged 10–20 years, Preston et al. (2014, 2017) reported a mean improvement of approximately 35% on untrained words after seven sessions with ultrasound visual feedback. Thus, treatment response varies across studies (and across individuals within studies), but there is evidence that /ɹ/ accuracy can improve after treatment that includes ultrasound visual feedback.

As most biofeedback programs for speech sound impairments are, to date, only available in a limited number of clinical settings, access to such programs remains a barrier. One study by Bernhardt et al. (2008) reported on a consultative model of treatment with ultrasound visual feedback for several children with /ɹ/ distortions whose accuracy was below 10% in words. They provided intensive training of up to 3–4 hr over the course of 2 days. Individual responses to the intensive consultative treatment ranged from an increase in accuracy of /ɹ/ of less than 10% to almost 90%, suggesting that short-term intensive treatment may be one feasible approach to enhance access.

Intensity of Service Delivery

The frequency with which clients are treated can make a difference in therapeutic outcomes for individuals with communication impairments. Providing one treatment session per week, as is the case in many schools (Brandel & Frome Loeb, 2011), may not be sufficient for individuals who have errors that have persisted for many years. This may be because of the need to replace an existing erred motor plan with a more accurate motor plan; even if acquisition of a target sound occurs during a treatment session, speech movements practiced successfully during therapy with clinician support may be quickly unlearned because of the many hours speaking without clinician support between therapy sessions. For example, training /ɹ/ in words to 80% accuracy in therapy once per week may result in poor generalization because of the many hours of nontherapeutic encounters and erred practice with /ɹ/.

A handful of studies have explored variations in service delivery for individuals with speech disorders, and it is generally the case that more frequent service results in greater improvement in speech sound accuracy (Kaipa & Peterson, 2016). For example, Allen (2013) reported that preschoolers with speech sound disorders made greater progress in their speech sound accuracy when treatment was provided three times a week than when it was provided once per week (with the total number of treatment hours held constant). Namasivayam et al. (2015) observed greater improvement in the speech sound accuracy of children with childhood apraxia of speech when they were treated twice per week rather than once per week. Similarly, intensive treatment programs appear to be more efficacious than less intense programs for treating speech sound errors in individuals with cleft palate (Albery & Enderby, 1984). For some adolescents with persisting speech errors associated with histories of childhood apraxia of speech, intensive treatment with ultrasound visual feedback has been found to facilitate acquisition of speech sounds that were in error, and it may enable generalization for some clients (Preston et al., 2016). Thus, an intensive treatment program might facilitate improved speech sound accuracy for adolescents and young adults with RSSEs, but limited data are available.

Auditory Speech Perception Training

Although obtaining proper articulator placement is essential for acceptable productions of speech sounds, auditory perception of speech is clearly an important component to acquiring new speech sounds. Speech perception has been found to differ in many children with speech sound disorders (Bird & Bishop, 1992; Cabbage, Hogan, & Carrell, 2016; Nijland, 2009; Ohde & Sharf, 1988; Preston et al., 2012; Rvachew & Jamieson, 1989; Rvachew, Ohberg, Grawburg, & Heyding, 2003; Shuster, 1998). Auditory perception training has been a component of speech therapy programs for decades (e.g., Van Riper, 1963), although it is not necessarily ubiquitously implemented. Attention to subtle acoustic phonetic detail may be particularly important for remediating /ɹ/ distortions given the unique acoustic feature of a low third formant. Thus, it may be necessary to explicitly instruct individuals with RSSEs to recognize phonetically acceptable and unacceptable productions of sounds in natural speech (Preston, Irwin, & Turcios, 2015). Rvachew, for example, has shown in several studies that auditory perceptual training with preschool and young school-age children with speech sound disorders enhances speech sound accuracy (Rvachew, 1994; Rvachew, Nowak, & Cloutier, 2004; Rvachew, Rafaat, & Martin, 1999), although structured auditory perceptual training programs for older clients do not appear to have been reported. In addition, because self-assessment of one's own speech sound accuracy may be poor for some individuals with speech sound disorders (Shuster, 1998), training clients to monitor and judge the accuracy of their speech has been shown to be clinically beneficial (Koegel, Koegel, & Ingham, 1986; Koegel, Koegel, Ingham, & Van Voy, 1988; Ruscello & Shelton, 1979; Shriberg & Kwiatkowski, 1987). Thus, explicitly requiring clients to listen to and evaluate the accuracy of the speech of others and of themselves may be an important component to facilitating speech sound learning.

Purpose and Hypotheses

The purpose of the present report was to evaluate whether observable changes in rhotic accuracy could be facilitated in adolescents and young adults with RSSEs using a theoretically motivated treatment program that included four key elements: a chaining-based practice and feedback schedule structured around PML, ultrasound biofeedback, intensive service delivery, and auditory perceptual training. Importantly, the goal was to evaluate whether changes can be observed in the speech of individuals with long-standing RSSEs after a program that is designed to include these elements (rather than to test each element independently). We hypothesized that measurable changes in speech sound acquisition and generalization can be achieved when the aforementioned elements are included.

Method

Participants

This case series includes four individuals who participated in an intensive 1-week program addressing rhotic distortions. All participants had histories of several years of speech therapy addressing rhotics, and all had rhotic errors that were evident at the word level and in conversation. As all participants lived geographically distant from the treatment site, a Skype session before the start of the study was used to confirm the presence of rhotic distortions in words and in conversation. Participants were of ages 13, 17, 21, and 22 years; all names reported in this case series are pseudonyms. This study was carried out in accordance with the recommendations of the Syracuse University Institutional Review Board. Written consent was provided by all participants or their guardians. In the case of the two minor participants, written assent was also obtained.

As all participants reportedly performed well academically, none had received speech therapy through their local education agencies in the past year. Liam was a 13-year-old who began speech and language therapy at the age of 3 years to address speech sound production. He reportedly had intermittent speech therapy throughout elementary and middle school and also received occasional private speech therapy; at the time of the study, he only had a distortion of /ɹ/. Katia was a 17-year-old who had normal speech development with the exception of her distorted /ɹ/ and had also received speech therapy twice per week during the ages of 13–15 years. Meredith was a 21-year-old college student who also had normal speech development with the exception of a distorted /ɹ/ and had received speech therapy at a university clinic and via private speech services over school breaks starting at the age of 19 years. James was a recent college graduate with a preexisting diagnosis of childhood apraxia of speech who had speech-language therapy from the age of 3 to 17 years, and he reported having infrequent speech sessions between college semesters. He continued to have occasional vowel distortions, epenthesis in consonant clusters, and infrequent distortions or substitutions of other phonemes, but his most apparent sound-specific error was on /ɹ/. He reported that he had difficulty perceiving errors in his speech.

Each participant attended a 1-week program consisting of approximately 2 hr in the morning and 2 hr in the afternoon for 5 days. The assessment and treatment schedule was similar for each participant and is outlined in the Appendix.

All participants passed a pure-tone hearing screening at 500, 1000, 2000, and 4000 Hz bilaterally at 20 dB. Participants also passed a vision screening using the Snellen eye chart. In addition, to characterize speech and language skills, several standardized and nonstandardized tests were administered, as outlined below. Participant information is presented in Table 1.

Assessments

Speech

Word-level speech sound accuracy was assessed with the Goldman-Fristoe Test of Articulation–Second Edition (Goldman & Fristoe, 2000). A multisyllabic word repetition task (Preston & Edwards, 2007) was administered to evaluate segmental accuracy and lexical stress in polysyllabic words. In addition, stimulability was evaluated through imitation of 11 syllables with /ɹ/, each repeated three times (cf. Miccio, 2002). Accuracy on /ɹ/ was assessed and tracked with generalization probes and a sentence imitation task described below.

A maximum performance task was administered to evaluate speech motor functioning (Thoonen, Maassen, Gabreels, & Schreuder, 1999; Thoonen, Maassen, Wit,

Table 1. Participant demographic information and results from standardized and nonstandardized assessments.

Measure	Value	Participant Liam	Katia	Meredith	James
Age	Years; months	13; 11	17; 4	21; 7	22; 8
GFTA-2	Standard score	70	69	74	< 40 (raw = 11)
	Percentile	2	< 1	< 1	< 1
Multisyllabic word repetition	Percent consonants correct	92	92	95	95
	Percent lexical stress correct	100	95	80	80
Stimulability probe	/ɹ/ Onset percent correct	0	0	67	0
	/ɹ/ Rhyme percent correct	0	0	57	92
Max performance task	Dysarthria score (out of 2)	0	0	0	0
	Apraxia score (out of 2)	2	0	0	2
PPVT-4	Standard score	116	107	96	120
CELF-5	Recalling Sentences scaled score	16	9	10	12 (raw = 73)
	Formulated Sentences scaled score	11	12	12	9 (raw = 43)
CTOPP-2	Elision scaled score	10	9	12	11
	Blending Words scaled score	13	13	11	11
	Phoneme Isolation scaled score	11	12	13	13
Nonword repetition	Percent consonants correct	96	98	94	92
WASI-II	Matrix reasoning *t* score	47	42	52	57

Note. At the time of the intensive treatment program, James's age was 22 years 8 months. This exceeded the maximum age in the standardization sample for both the GFTA-2 and the CELF-5 (21;11 for both tests). For comparison purposes, James was compared with were maximum age band on both the GFTA-2 and the CELF-5; however, raw score data are also provided. Standard scores have a mean of 100 and a standard deviation of 15, scaled scores have a mean of 10 and a standard deviation of 3, and *t* scores have a mean of 50 and a standard deviation of 10. GFTA-2 = Goldman-Fristoe Test of Articulation–Second Edition; PPVT-IV = Peabody Picture Vocabulary Test–Fourth Edition; CELF-5 = Clinical Evaluation of Language Fundamentals–Fifth Edition; CTOPP-2 = Comprehensive Test of Phonological Processing–Second Edition; WASI-II = Wechsler Abbreviated Scales of Intelligence–Second Edition.

Gabreëls, & Schreuder, 1996). Participants produced sustained phonemes /ɑ/, /f/, /s/, and /z/ as well as repeated productions of /mɑmɑ/, /pʌ/, /tʌ/, /kʌ/, and /pʌtʌkʌ/. From the maximum durations and rate of sound sequences on this task, scores were derived for both dysarthria and apraxia characteristics, with 0 reflecting *no dysarthria/apraxia,* 1 representing *undetermined,* and 2 representing *dysarthria/apraxia* (Rvachew, Hodge, & Ohberg, 2005).

Additional Assessments

In addition, oral language, phonological processing, and nonverbal assessments were administered to provide descriptive information. This information was collected to aid in the interpretation of individual factors that may relate to treatment responses.

To characterize oral language abilities, the Peabody Picture Vocabulary Test–Fourth Edition (Dunn & Dunn, 2007) and the Recalling Sentences and Formulated Sentences of the Clinical Evaluation of Language Fundamentals–Fifth Edition (Wiig, Semel, & Secord, 2013) were administered.

From the Comprehensive Test of Phonological Processing–Second Edition (Wagner, Torgesen, Rashotte, & Pearson, 2013), the Elision, Blending Words, and Phoneme Isolation subtests were administered. In addition, from the nonword repetition task (Dollaghan & Campbell, 1998), percent phonemes correct was computed.

The Matrix Reasoning subtest of the Wechsler Abbreviated Scales of Intelligence–Second Edition (Wechsler, 2011) was administered to characterize visual reasoning abilities.

Treatment Targets and Progress Monitoring

All participants were treated on rhotics in onset (singletons and clusters) and rhyme (nucleus and coda) positions. Progress was tracked through generalization probes (Supplemental Material S6), word lists read by the participant with no model. Probes consisted of untreated words with /ɹ/ as onset singletons (e.g., "radio," 25 items), onset clusters (e.g., "princess," 50 items), coda (e.g., "bookstore," 25 items), and the nucleus /ɝ/ (e.g., "bird," 25 items). Participants completed home recordings of the probe lists 1 week before the start of the study using their smartphones and again at the treatment site on Days 1 (before treatment) and 3 (partway through treatment) and twice on Day 5 (after treatment) of the study week; they also completed a recording of the probes at home 3 weeks after treatment, which was sent to the researchers to evaluate retention. Probes completed via smartphone were recorded using a voice memo or video application native to the device. These recordings were collected to obtain additional speech samples because of the limited duration of the treatment. All audio tracks were converted to .wav files in Audacity (Version 2.0.3, Audacity Team, 2008). Sampling rates for audio tracks submitted from smartphones were a minimum of 44 kHz. Recordings at the treatment site were conducted with a Sennheiser MKE-2 lapel microphone and were recorded at 44 kHz into Praat (Version 5.3.82; Boersma & Weeninck, 2015).

In addition to the word-level generalization probes, participants completed a sentence imitation task on Day 1 (before treatment), Day 3 (partway through treatment), and Day 5 (after treatment) of the study week. The sentence imitation task consisted of 15 sentences containing /ɹ/ in various contexts (e.g., "George stayed true to his roots."). Contexts assessed in the sentence imitation task included onset singleton (e.g., "rowing," 10 items), onset cluster (e.g., "true," 14 items), nucleus (e.g., "minor," 12 items), coda singleton (e.g., "far," nine items), and coda cluster (e.g., "smart," eight items).

Treatment Overview

Assessment and treatment were provided by two American Speech-Language-Hearing Association–certified SLPs who were familiar with the treatment procedures (the two authors), with each SLP treating each participant for half of the therapy (7 hr per SLP). The first morning included the assessments described above followed by a brief overview of the ultrasound, identification of different parts of the tongue, and a description of the participant's tongue shape during production of distorted /ɹ/. Comparisons were made to some of the key features of a correct /ɹ/ such as elevation of the tongue tip or blade, lowering of the tongue dorsum, tongue root retraction (in sagittal view), and elevation of the lateral margins of the tongue with midline grooving (in coronal view). During each hour of treatment, either onset (which included both singletons and clusters) or rhyme (which included both nucleus and coda) was targeted.

Auditory Training

The breakdown of treatment tasks during the sessions is provided in Table 2. The first seven sessions included 6–8 min of auditory perceptual judgments of recordings of correct and distorted /ɹ/, which came from individuals with and without speech sound errors. The task was administered through a researcher-developed web application and was generally modeled after the Speech Assessment and Interactive Learning System (Rvachew, 1994) but with a wider range of stimuli. Each 50-item module presented /ɹ/ in either onset or rhyme. Within each module, 40%–60% of items were correct and 40%–60% were incorrect (i.e., coming from individuals with speech sound disorders). Participants listened to each recorded word while the printed word appeared on the screen (e.g., "rat") and clicked "right" or "wrong" to evaluate the accuracy of the rhotic in the word production. Participants could replay items up to five times before making a decision. As shown in Table 2, auditory perceptual training was not included in the last seven sessions but was replaced with randomized production practice (see below).

Table 2. Overview of each hour-long treatment session.

Sessions 1–7	Sessions 8–14
Auditory perceptual judgments (~6–8 min)	—
Time period A: with ultrasound (12 min)	Time period A: with ultrasound (12 min)
Time period B: without ultrasound (12 min)	Time period B: without ultrasound (12 min)
Time period C: with ultrasound (12 min)	Time period C: with ultrasound (12 min)
Time period D: without ultrasound (12 min)	Time period D: without ultrasound (12 min)
—	Randomized practice (6–8 min)

Production Training

The use of the ultrasound was time dependent. Each session included four 12-min periods, A, B, C, and D, which were controlled by the use of a timer. Periods A and C always included the use of ultrasound visual feedback during production practice using an Echo Blaster 128 ultrasound with a PV 6.5 transducer (Telemed, UAB). Periods B and D included no ultrasound (to facilitate generalization and reduce reliance on the feedback).

In addition to the time-dependent structure, there were two phases of treatment to each hour of treatment. The first phase, prepractice, involved eliciting 12 correct syllable-level targets. For example, when /ɹ/ onset was treated, the participant had to produce three tokens of two onset singletons and two onset clusters (e.g., /ɹi/, /ɹo/, /bɹʌ/, /θɹʌ/). To elicit correct versions of these targets, the clinician was free to use phonetic placement cues, verbal descriptions of articulatory movement, references to images of the articulatory movements, and shaping techniques (e.g., shaping /ɹ/ from /ɑ/ or /l/). As the prepractice always began in Period A, prepractice also included reference to the ultrasound images. Thus, the SLP coached the client to modify the tongue shape (e.g., in sagittal view, this included a focus on raising the anterior tongue, lowering the dorsum, and retracting the tongue root). Verbal descriptions and visual cues were provided, such as drawing a tongue shape on a transparency laid over the screen for the client to try to copy. Once a tongue shape was identified that resulted in a perceptually correct form of the target, this was traced over the screen and discussed with the client so they had a clear understanding of a tongue shape that facilitated accurate /ɹ/. Participants Katia, Meredith, and James practiced primarily with sagittal views, as the cues for tongue root retraction and dorsum lowering were the most facilitative; however, Liam practiced primarily with a coronal view as the focus on elevation of the lateral margins of his tongue was judged by the SLPs to be most facilitative. An example of prepractice is included in Supplemental Material S1.

Once the prepractice criterion of 12 correct productions was met for that session (which took most of the 48 min of practice in the early sessions and only 1–2 min in later sessions), prepractice ended, and structured chaining practice began. Structured chaining practice included a predetermined series of blocks of six consecutive attempts at various linguistic targets. Chaining procedures always began at the syllable level and progressed to monosyllabic words, multisyllabic words, set phrases, and self-generated sentences based on the participant's performance. Each higher level of linguistic practice included the target syllable (e.g., /ɹo/, *roll, rolling, rolling the ball,* and a self-generated sentence with *rolling*). If the participant achieved at least five of six correct trials in a block, she or he progressed to the next linguistic level of the chain. If, at any time, the participant did not achieve five of six correct trials, the next block of six trials began a new chain starting at the syllable level (e.g., /bɹʌ/, *broom, broomstick, witch on a broomstick,* and a self-generated sentence with *broomstick*). If the participant successfully completed a chain (i.e., produced the self-generated sentence accurately) twice in a session, the chain was discontinued from treatment in the subsequent session and replaced with a new chain with the same target syllable. During half of the sessions, variable practice was implemented by encouraging participants to produce target words, phrases, and sentences with variations in prosody (e.g., practice *rolling* as a question, as a command, slowly, quickly, or loudly; cf. Preston et al., 2014).

In addition, feedback was adaptive during structured chaining practice. At the syllable level, participants received verbal feedback on five of six trials (four trials with KP and KR feedback, one with only KR feedback, and one with no feedback). However, feedback was reduced in frequency and transitioned to less verbal KP feedback at subsequent linguistic levels such that, at the sentence level, the participant received verbal feedback on only three of six trials (one with verbal KP feedback, two with verbal KR feedback, and three with no feedback). In addition, practice variability was included in half of the sessions by encouraging attempts of target words, phrases, and sentences with variations in prosody (e.g., practicing *broomstick* slow, fast, as a question, as a command, loud, and neutral). A sample datasheet with the feedback schedule and sample stimuli can be found in Supplemental Material S5.

The same structured chaining practice procedures were implemented in Periods B and D, with two exceptions: Ultrasound visual feedback was not used, and participants were required to self-evaluate the auditory accuracy of their productions on 50% of the trials. A sample video of structured chaining practice with and without the ultrasound is included in Supplemental Materials S2 and S3.

Randomized Practice

During the final seven sessions, the last 6–8 min of the session were spent in randomized practice. Targets at the highest level of the chain (i.e., at least five of six correct) during Periods A–D of that session were chosen. During randomized practice, the order of stimuli was mixed, and a dice roll was used to determine the number of trials on the target word or phrase (one to six). The SLP provided only delayed KR feedback (i.e., a 2- to 3-s pause followed by "good" or "not quite") on every other attempt (for 50% feedback).

Data Analysis and Reliability

Each recording of the word-level generalization probes was segmented into individual words. All recorded words were then randomized and grouped into modules of 100 for ratings that were completed by four research assistants.[1] The four listeners independently rated all tokens as correct or incorrect (using a script written in Praat) while blind to whether each recorded word was collected before, during, or after treatment. Fleiss' kappa, an estimate of reliability across multiple raters, was computed in R Version 3.2.2 (R Core Team, 2015) and was estimated at .59 (95% confidence interval [CI] [0.58, 0.60]). The final data represent the average of the four listeners.

In addition to the word-level generalization probes, learning was assessed by a sentence imitation task. Ratings on each production of /ɹ/ were obtained from two research assistants who had been trained in scoring procedures for the sentence imitation task. The Fleiss' κ across both raters was .71 (95% CI [0.63, 0.79]). The final data represent the average of both listeners.

Percent change in /ɹ/ accuracy is reported for both word- and sentence-level generalization measures. In addition, d_2, a standardized measure of change (mean difference from pretreatment to posttreatment divided by the pooled standard deviation), is also reported (Beeson & Robey, 2006).

Four sessions per participant (16 sessions in total) were reviewed by a research assistant to judge fidelity to the treatment protocol with regard to feedback (appropriate rate and type of feedback given) and provision of self-judgment opportunities. Because each SLP delivered therapy for half of the 14 hr, the four sessions per participant were composed of two from each treating SLP. With regard to feedback, the specified verbal feedback (KP + KR, KR only, or no feedback) was provided, on average, 97.9% of the time in sessions that were reviewed (*SD* = 1.9%, range = 94.1%–100%). Participants were asked to self-judge 50% of their productions, before SLP feedback, during portions of the therapy. The SLPs prompted participants to self-judge, on average, 98.7% of the time (*SD* = 1.8%, range = 95.5%–100%).

In addition to treatment fidelity, interrater reliability to SLP judgment of acoustic correctness was assessed because the treatment relies on the treating clinician's immediate decisions about acoustic acceptability of the participant's productions. When reviewing the 16 sessions, the research assistant calculated agreement with the treating clinician's judgment of accuracy (verbal KR) during structured practice only when KR or KR + KP feedback was given (trials without feedback were not calculated in the overall percent agreement as no KR feedback was provided to the participant). The research assistant agreed with the treating clinician's KR decision about acoustic accuracy, on average, 97% of the time in the sessions that were reviewed (*SD* = 1.2%, range = 95.4%–98.6%).

[1]The four research assistants were graduate and undergraduate students who were native speakers of rhotic dialects of American English. They had taken coursework in phonetics and speech sound disorders. All were aware of the response definitions of correct and incorrect and had been trained to at least 90% agreement with expert listeners on 200 recorded tokens judging accuracy of /ɹ/.

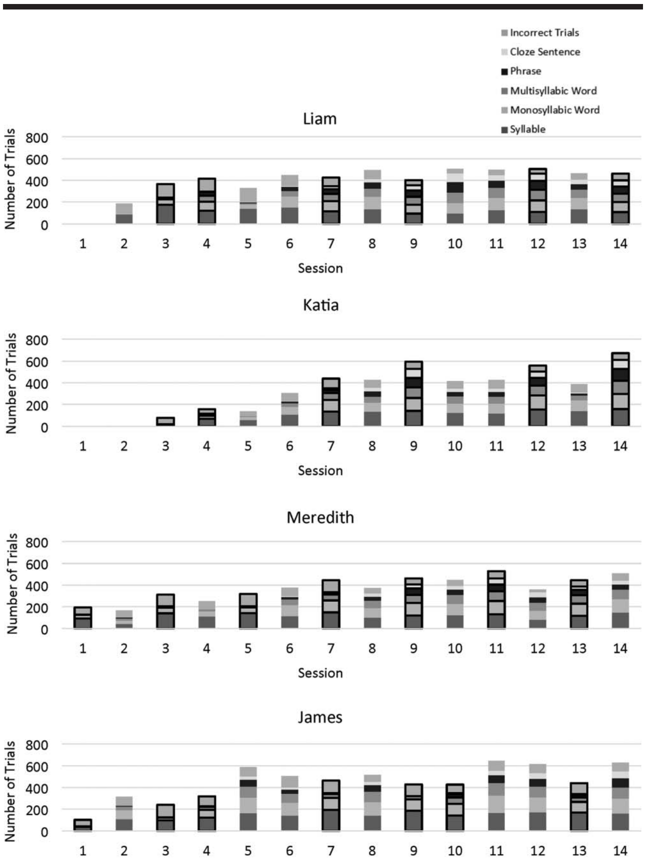
Results

Within-Session Performance

Within-session performance was tracked by the number of correct trials per hour of treatment. As can be seen in Figure 1, the first few sessions for each participant showed fewer correct trials (and fewer overall attempts) than in later sessions. It is also evident that, over the course of the week, more practice was occurring at higher levels of complexity, suggesting that the clinician perceived greater speech sound accuracy. The total number of practice trials in structured chaining practice in the 1-week program was 4,475 for Liam, 4,596 for Katia, 5,195 for Meredith, and 6,219 for James; this resulted in an average of over 300 structured practice trials per hour-long session for each participant.

For the first seven sessions, participants completed auditory perception training modules in which they were required to listen to recordings and judge the accuracy of recorded productions. Participants' agreement with the researchers' judgments of those tokens were as follows: Liam, 93%; Katia, 95%; Meredith, 88%; and James, 85%.

Figure 1. Acquisition data (trials scored correct/incorrect within the session) by the treating clinician in structured chaining practice. Columns outlined in black represent sessions where /ɹ/ was treated in the onset position. Columns with no outline represent sessions where /ɹ/ was treated in the rhyme position.



In addition, during structured chaining practice, participants were queried about the accuracy of their productions during Periods B and D, and these ratings were compared against the treating SLP's judgment. Over the course of the study, participants' self-ratings were in agreement with those of the SLPs: Liam, 87%; Katia, 91%; Meredith, 85%; and James, 84%.

Generalization

The primary outcome of interest was participants' performance on the word-level generalization probes and the sentence imitation task. These scores are presented by participant in Figure 2. Samples of pretreatment and posttreatment generalization probes are available in Supplemental Material S4. The difference in the proportion of generalization probe words judged correct from pretreatment to posttreatment was computed for each participant (Unicomb, Colyvas, Harrison, & Hewat, 2015); the 95% CI surrounding the difference score was computed using the software package R Version 3.2.2 (R Core Team, 2015). These difference scores are shown in Figure 3 and suggest that the change from pretreatment to posttreatment was reliably above zero for all four participants. In addition, a chi-square (χ^2) analysis was used to compare the proportion of generalization probe items rated as correct before and after therapy for each participant in IBM SPSS Statistics 24.

Liam

Liam's pretreatment accuracy for /ɹ/ in the onset position was 17%, which increased to 90% posttreatment (d_2 = 16.3). His accuracy for /ɹ/ in the rhyme position increased from 4% to 93% (d_2 = 24.5). There was a significantly greater proportion of Liam's probe words rated as "correct" during posttreatment probes than during pretreatment probes, $\chi^2(1) = 1882$, $p < .001$. In addition, his accuracy of /ɹ/ in sentences increased from 6% pretreatment to 81% posttreatment.

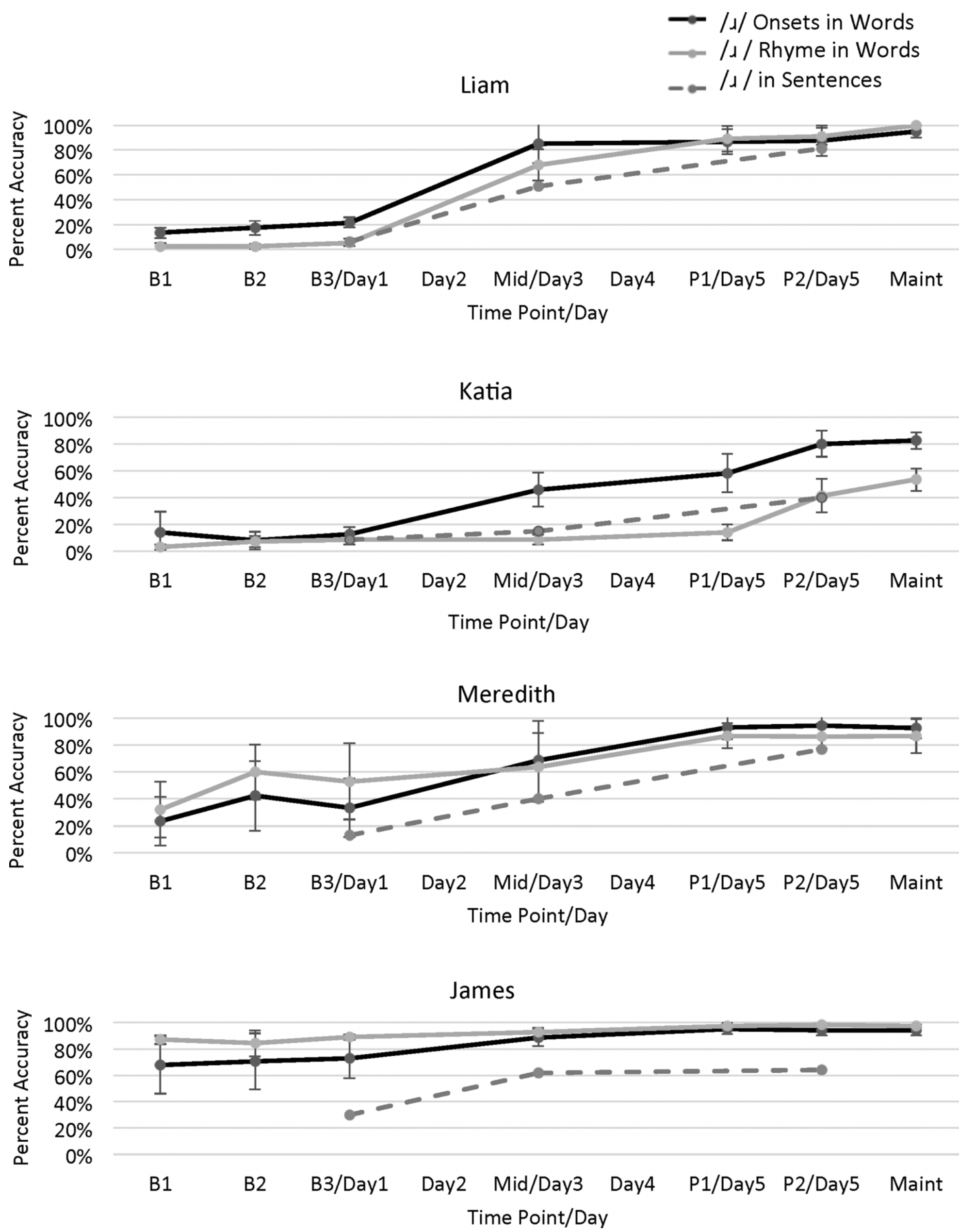
Katia

Katia showed a strong response to treatment for /ɹ/ in the onset position, showing a 62% increase (d_2 = 7.5) on generalization probes (from 12% at baseline to 74% posttreatment). An increase of 30% was observed on generalization probes for /ɹ/ rhyme (from 6% to 36%, d_2 = 2.6). A significantly greater proportion of probe words was rated as "correct" after treatment than before treatment, $\chi^2(1) = 812$, $p < .001$. In addition, her sentence level accuracy for /ɹ/ also increased from 8% to 40%.

Meredith

Meredith's pretreatment accuracy for /ɹ/ in the onset position was 33%, which increased to 94% posttreatment (d_2 = 11.6). Her accuracy for /ɹ/ in the rhyme position showed a 39% increase (from 48% to 87%, d_2 = 5.2). There was a statistically significant difference between the proportion of probe words rated as "correct" after therapy

Figure 2. Performance on word-level generalization probes and sentences. Percent accuracy for word-level probes is the average rating of all generalization probe words by four listeners. /ɹ/ in the onset position of words is shown by solid black lines. /ɹ/ in the rhyme position of words is shown by gray lines. /ɹ/ across various word positions at the sentence level is shown by dotted lines. B = baseline; Mid = midpoint of treatment; P = post-treatment; Maint = maintenance 3 weeks after treatment.



compared with before therapy, $\chi^2(1) = 880$, $p < .001$. Her sentence level accuracy for /ɹ/ increased from 13% to 77%.

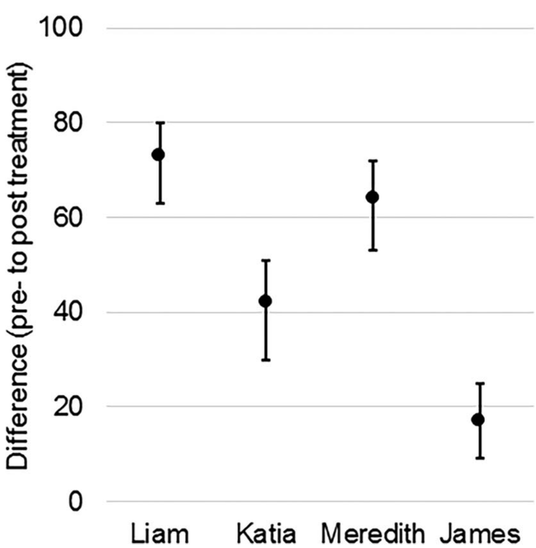
James

Before treatment, James was relatively more successful than the other participants at the word level, producing 70% correct /ɹ/ in the onset position and 86% in the rhyme position. After treatment, his accuracy increased to 95% for onset (d_2 = 15.7) and 98% for rhyme (d_2 = 7.2). There was a significantly greater proportion of /ɹ/ words rated as "correct" posttreatment than pretreatment, $\chi^2(1) = 233$, $p < .001$. Moreover, his sentence level accuracy for /ɹ/ increased from 30% pretreatment to 64% posttreatment.

Group Level Changes

Averaged across all word positions, probe administrations, and participants, the mean pretreatment accuracy on

Figure 3. Difference scores on word-level generalization probes with 95% confidence intervals. Circles reflect the difference in proportion of words rated correct from pretreatment to posttreatment. None of the confidence intervals encompass 0, suggesting statistically significant change in rhotic accuracy on the generalization probes (Unicomb et al., 2015).



the word-level generalization probes was 35% (SD = 33%), and the mean posttreatment accuracy was 83% (SD = 19%). This represents an overall increase of 1.9 standard deviations. In addition, as shown in Figure 3, the difference between pretreatment and posttreatment was reliably greater than zero. On the sentence imitation task, there was an increase from 14% (SD = 11%) to 66% (SD = 18%), an increase of 3.5 standard deviations.

Discussion

This case series investigated whether an intensive 1-week program could result in acquisition and generalization for individuals with RSSEs that involved rhotic distortions. The results of this investigation confirm the hypothesis that it is possible to effect change in individuals with these long-standing errors with a treatment program that incorporates several PMLs, ultrasound visual feedback, an intensive treatment schedule, and an auditory perception training. There was a substantial increase in accuracy at both the word and sentence levels across all four participants. Change scores from pretreatment to posttreatment were reliably greater than 0, averaging approximately 50% above baseline, which is roughly similar in magnitude to previous reports exploring different treatment schedules (e.g., Adler-Bock et al., 2007; Bacsfalvi & Bernhardt, 2011; McAllister Byun et al., 2014). Rhotics in both onsets and rhymes increased, although, for Katia, greater improvement was observed for onset than able improvements at the sentence level for all participants. Follow-up data 3 weeks after the treatment suggested that gains had been maintained or that continued improvement had occurred (see Figure 2). Therefore, one important outcome of this report is evidence that, despite habituating these errors for many years and responding poorly to prior speech strategies, the individuals with RSSEs treated here improved their speech sound accuracy and showed evidence of motor learning. Thus, we assert that dismissing individuals with RSSEs from speech-language therapy services due to "failure to make progress" should be considered only once alternatives to traditional therapy methods have been explored.

It is possible that the intensity of the treatment was an important factor that facilitated change. Traditional models of treatment that involve one or two sessions per week allow for many opportunities for negative practice between sessions (i.e., returning to the old, erred articulatory movements). Thus, intensive treatment may be an important component of treatment for "undoing" previously established sound errors and guarding against regression between treatment sessions. Intensive treatment may have the advantage of offering a sufficiently high dosage of treatment to unlearn an erred motor pattern and replace it with a new pattern. Because these individuals were between 13 and 22 years old, they could tolerate drill practice and achieved over 4,400 total practice trials in the course of 1 week (mean per session = 366, SD = 165). The traditional approach of distributed practice (e.g., one to two sessions per week) may result in slow generalization because of the many hours of nontherapeutic encounters and therefore erred practice with speech targets between sessions. Thus, limiting the time between practice sessions may be important to consider when training speech in individuals with RSSEs (cf. Allen, 2013; Kaipa & Peterson, 2016).

As previous studies have shown some success with biofeedback and a motor learning approach to treatment, it is possible that the biofeedback played an important role as well. Previous studies have shown that treatment programs that include ultrasound can lead to improved speech sound accuracy in adolescents and adults with RSSEs (e.g., Adler-Bock et al., 2007; Bacsfalvi & Bernhardt, 2011; Bernhardt et al., 2008; Cleland et al., 2015; Modha et al., 2008; Preston et al., 2014, 2017). All participants reported anecdotally that visualizing their tongue allowed them to better understand the articulatory targets and the nature of their errors. However, each of the participants also reported that the ultrasound was particularly helpful during the first 2 or 3 days of treatment and that, during the final days, they relied less on the visual feedback. Thus, fading the use of the visual feedback based on client performance may be preferred.

Finally, the auditory training may also have played a role in treatment outcomes. In general, the participants were all successful at identifying correct and distorted /ɹ/ in recordings of the speech of others (≥ 85%), and all showed good agreement with the clinician's judgment of their errors during treatment (≥ 84%). Requiring participants to be accountable for self-monitoring, and comparing their judgments against the SLP's judgments, may have facilitated their need to attend to subtle distinctions in their productions (Koegel et al., 1986; Ruscello & Shelton, 1979). Although experimental studies have revealed differences in speech perception between children with and without RSSEs (Cabbage et al., 2016; Shuster, 1998), there are currently no clinically validated procedures to reliably detect speech perception problems in this age group (Preston et al., 2015); thus, it is unclear whether there were overt perceptual deficits in these participants and whether the perceptual training played a crucial role.

Caveats and Limitations

Practical limitations of this approach are evident. For example, biofeedback training requires access to technologies that may be expensive and also require training for the SLP. Intensive treatment programs require time commitments for both clients and clinicians, and third-party reimbursement systems are currently not well structured to support them. Auditory perceptual training, particularly for older clients with RSSEs, still requires development, validation, and commercialization. Thus, the most accessible element to the program described herein is likely the structured chaining practice, which only requires familiarity with theoretical principles and their structured implementation in therapy. On the basis of only this case series, however, it cannot be determined which of the elements of this treatment program are the essential ingredients. Notwithstanding these limitations, it is clinically important to note that rapid change in speech sound accuracy can be achieved under the conditions described here, even for individuals whose RSSEs have been seemingly intractable.

Moreover, the progress observed by the participants was objectively measurable by blinded listeners at both the word and sentence levels, but the results fall short of an ideal outcome (e.g., 100% accurate productions for all participants in words and sentences) at the conclusion of 1 week. As can be seen in Figure 2, there was generally a continued trend for improved productions at the 3-week follow-up suggesting that continued learning may have occurred once treatment was withdrawn. However, to reach the point of nearly 100% accurate production, continued treatment may be necessary.

The design of this case report involved progress monitoring from pretreatment to posttreatment and therefore did not involve a no-treatment comparison condition. However, several previous studies have shown stable performance on similar generalization probes over the course of 1–2 weeks for individuals with RSSEs (McAllister Byun & Hitchcock, 2012; McAllister Byun et al., 2014; Preston et al., 2014, 2017), and the long-standing nature of these errors suggests that it is not plausible that the improvements observed here are due simply to maturation effects over 1 week; thus, the improvements are likely attributable to the effects of the treatment. Whether other treatment approaches could provide similar outcomes remains an open question. Furthermore, replicating these results under more controlled experimental conditions (e.g., randomized controlled trials) and extending the results to additional populations in traditional settings (e.g., local educational agencies) would be of clinical value. In addition, as the primary aim was to maximize outcomes rather than to test the effects of a single variable, the observed changes may, at present, be attributed to the treatment "package" as a whole rather than to any specific element of the treatment.

One observable side effect of the treatment was vocal fatigue. Presumably because of the intensity of the treatment and the high number of practice trials in a 1-week period, temporary vocal fry was observed, particularly for Katia. Follow-up contact with this participant, as well as recordings submitted to the researchers 3 weeks after the intensive program, indicated that the vocal fry had subsided. Thus, this side effect was temporary but should be anticipated and managed in future attempts to replicate these results.

Finally, it is unlikely that all individuals with RSSEs are equally good candidates for this type of treatment program. The participants in this study all had age-appropriate cognitive and linguistic skills, adequate motivation, and the ability to sustain attention during a drill-oriented treatment. It is conceivable that individuals with cognitive or attention deficits, or individuals who are younger or less motivated, might respond differently.

Summary and Conclusions

This report provides evidence that speech sound acquisition and generalization can be achieved in an intensive motor-based treatment program that includes biofeedback and auditory training. The study provides evidence that at least some adolescents and adults with RSSEs have sufficient neuro-phonetic plasticity to achieve changes in speech sound production that generalize to untrained words and that improvements can be retained for at least 3 weeks. Clinically, the results show promise for this type of program, suggesting that at least some individuals with unresolved RSSEs can improve their speech sound accuracy with modifications to traditional therapy methods.

Acknowledgments

This study was supported by NIH Grant R03DC013152, which was granted to Jonathan L. Preston, PI. Thanks to Nina Chu, Kara Comins, Kerry McNamara, Lauren Peduto, Stephanie Urlass, Monique Uy, and Haley Weaver for their assistance with this project.

References

Adler-Bock, M., Bernhardt, B., Gick, B., & Bacsfalvi, P. (2007). The use of ultrasound in remediation of North American English /r/ in 2 adolescents. *American Journal of Speech-Language Pathology, 16*(2), 128–139. https://doi.org/10.1044/1058-0360(2007/017)

Albery, L., & Enderby, P. (1984). Intensive speech therapy for cleft palate children. *British Journal of Disorders of Communication, 19*(2), 115–124. https://doi.org/10.3109/13682828409007182

Allen, M. M. (2013). Intervention efficacy and intensity for children with speech sound disorder. *Journal of Speech, Language,* *and Hearing Research, 56*(3), 865–877. https://doi.org/10.1044/1092-4388(2012/11-0076)

Audacity Team. (2008). Audacity (Version 2.0.3) [Computer program]. Retrieved from http://audacityteam.org/

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

Bacsfalvi, P., & Bernhardt, B. M. (2011). Long-term outcomes of speech therapy for seven adolescents with visual feedback technologies: Ultrasound and electropalatography. *Clinical Linguistics & Phonetics, 25*(11–12), 1034–1043.

Beeson, P. M., & Robey, R. R. (2006). Evaluating single-subject treatment research: Lessons learned from the aphasia literature. *Neuropsychology Review, 16*(4), 161–169.

Bernhardt, B., Bacsfalvi, P., Adler-Bock, M., Shimizu, R., Cheney, A., Giesbrecht, N., . . . Radanov, B. (2008). Ultrasound as visual feedback in speech habilitation: Exploring consultative use in rural British Columbia, Canada. *Clinical Linguistics & Phonetics, 22*(2), 149–162.

Bird, J., & Bishop, D. V. M. (1992). Perception and awareness of phonemes in phonologically impaired children. *European Journal of Disorders of Communication, 27,* 289–311.

Boersma, P., & Weeninck, D. (2015). *Praat: Doing phonetics by computer* (version 5.3.82) [Computer software]. Retrieved from http://www.praat.org

Boyce, S. E. (2015). The articulatory phonetics of /r/ for residual speech errors. *Seminars in Speech and Language, 36*(4), 257–270.

Brandel, J., & Frome Loeb, D. (2011). Program intensity and service delivery models in the schools: SLP survey results. *Language, Speech, and Hearing Services in Schools, 42*(4), 461–490. https://doi.org/10.1044/0161-1461(2011/10-0019)

Cabbage, K. L., Hogan, T. P., & Carrell, T. D. (2016). Speech perception differences in children with dyslexia and persistent speech delay. *Speech Communication, 82,* 14–25.

Cleland, J., Scobbie, J. M., & Wrench, A. A. (2015). Using ultrasound visual biofeedback to treat persistent primary speech sound disorders. *Clinical Linguistics & Phonetics, 29*(8–10), 575–597.

Dollaghan, C. A., & Campbell, T. F. (1998). Nonword repetition and child language impairment. *Journal of Speech, Language, and Hearing Research, 41,* 1136–1146.

Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test–Fourth Edition*. Minneapolis, MN: Pearson.

Fabus, R., Raphael, L., Gatzonis, S., Dondorf, K., Giardina, K., Cron, S., & Badke, B. (2015). Preliminary case studies investigating the use of electropalatography (EPG) manufactured by Completespeech® as a biofeedback tool in intervention. *International Journal of Linguistics and Communication, 3*(1), 11–23.

Flipsen, P. (2015). Emergence and prevalence of persistent and residual speech errors. *Seminars in Speech and Language, 36*(4), 217–223.

Gibbon, F., & Lee, A. (2015). Electropalatography for older children and adults with residual speech errors. *Seminars in Speech and Language, 36*(4), 271–282.

Goldman, R., & Fristoe, M. (2000). *Goldman-Fristoe Test of Articulation–Second Edition*. Circle Pines, MN: AGS.

Guadagnoli, M. A., & Lee, T. D. (2004). Challenge point: A framework for conceptualizing the effects of various practice conditions in motor learning. *Journal of Motor Behavior, 36*(2), 212–224.

Hitchcock, E. R., & McAllister Byun, T. M. (2015). Enhancing generalisation in biofeedback intervention using the challenge point framework: A case study. *Clinical Linguistics & Phonetics, 29*(1), 59–75. https://doi.org/10.3109/02699206.2014.956232

Hitchcock, E. R., Harel, D., & McAllister Byun, T. M. (2015). Social, emotional, and academic impact of residual speech errors in school-aged children: A survey study. *Seminars in Speech and Language, 36*(04), 283–294. https://doi.org/10.1055/s-0035-1562911

Huang, H., Wolf, S. L., & He, J. (2006). Recent developments in biofeedback for neuromotor rehabilitation. *Journal of Neuroengineering Rehabilitation, 3,* 11. https://doi.org/1743-0003-3-11 [pii]10.1186/1743-0003-3-11

Kaipa, R., & Peterson, A. M. (2016). A systematic review of treatment intensity in speech disorders. *International Journal of Speech-Language Pathology, 18*(6), 507–520. https://doi.org/10.3109/17549507.2015.1126640

Klein, H. B., McAllister Byun, T., Davidson, L., & Grigos, M. I. (2013). A multidimensional investigation of children's /r/ productions: Perceptual, ultrasound, and acoustic measures. *American Journal of Speech-Language Pathology, 22*(3), 540–553. https://doi.org/10.1044/1058-0360(2013/12-0137)

Koegel, L. K., Koegel, R. L., & Ingham, J. C. (1986). Programming rapid generalization of correct articulation through self-monitoring procedures. *Journal of Speech and Hearing Disorders, 51*(1), 24–32.

Koegel, R. L., Koegel, L. K., Ingham, J. C., & Van Voy, K. (1988). Within-clinic versus outside-of-clinic self-monitoring of articulation to promote generalization. *Journal of Speech and Hearing Disorders, 53*(4), 392–399.

Maas, E., Robin, D. A., Austermann Hula, S. N., Freedman, S. E., Wulf, G., Ballard, K. J., & Schmidt, R. A. (2008). Principles of motor learning in treatment of motor speech disorders. *American Journal of Speech-Language Pathology, 17*(3), 277–298. https://doi.org/10.1044/1058-0360(2008/025)

McAllister Byun, T. M., & Campbell, H. (2016). Differential effects of visual-acoustic biofeedback intervention for residual speech errors. *Frontiers in Human Neuroscience, 10,* 567. https://doi.org/10.3389/fnhum.2016.00567

McAllister Byun, T. M., & Hitchcock, E. R. (2012). Investigating the use of traditional and spectral biofeedback approaches to intervention for /r/ misarticulation. *American Journal of Speech-Language Pathology, 21*(3), 207–221. https://doi.org/10.1044/1058-0360(2012/11-0083)

McAllister Byun, T. M., Hitchcock, E. R., & Swartz, M. T. (2014). Retroflex versus bunched in treatment for rhotic misarticulation: Evidence from ultrasound biofeedback intervention. *Journal of Speech, Language, and Hearing Research, 57*(6), 2116–2130.

McAllister Byun, T. M., Swartz, M. T., Halpin, P. F., Szeredi, D., & Maas, E. (2016). Direction of attentional focus in biofeedback treatment for /r/ misarticulation. *International Journal of Language & Communication Disorders, 51*(4), 384–401. https://doi.org/10.1111/1460-6984.12215

McCormack, J., McLeod, S., McAllister, L., & Harrison, L. J. (2009). A systematic review of the association between childhood speech impairment and participation across the lifespan. *International Journal of Speech-Language Pathology, 11*(2), 155–170.

Miccio, A. W. (2002). Clinical problem solving: Assessment of phonological disorders. *American Journal of Speech-Language Pathology, 11*(3), 221–229.

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

Namasivayam, A. K., Pukonen, M., Goshulak, D., Hard, J., Rudzicz, F., Rietveld, T., . . . van Lieshout, P. (2015). Treatment intensity and childhood apraxia of speech. *International*

Journal of Language & Communication Disorders, 50(4), 529–546. https://doi.org/10.1111/1460-6984.12154
Newell, K., Carlton, M., & Antoniou, A. (1990). The interaction of criterion and feedback information in learning a drawing task. *Journal of Motor Behavior, 22*(4), 536–552.
Nijland, L. (2009). Speech perception in children with speech output disorders. *Clinical Linguistics & Phonetics, 23*(3), 222–239.
Ohde, R. N., & Sharf, D. J. (1988). Perceptual categorization and consistency of synthesized /r-w/ continua by adults, normal children and /r/-misarticulating children. *Journal of Speech and Hearing Research, 31,* 556–568.
Preston, J. L., & Edwards, M. L. (2007). Phonological processing skills of adolescents with residual speech sound errors. *Language, Speech, and Hearing Services in Schools, 38,* 297–308.
Preston, J. L., Felsenfeld, S., Frost, S. J., Mencl, W. E., Fulbright, R. K., Grigorenko, E. L., . . . Pugh, K. R. (2012). Functional brain activation differences in school-age children with speech sound errors: Speech and print processing. *Journal of Speech, Language, and Hearing Research, 55*(4), 1068–1082. https://doi.org/10.1044/1092-4388(2011/11-0056)
Preston, J. L., Irwin, J. R., & Turcios, J. (2015). Perception of speech sounds in school-aged children with speech sound disorders. *Seminars in Speech and Language, 36*(4), 224–233.
Preston, J. L., Leece, M. C., & Maas, E. (2016). Intensive treatment with ultrasound visual feedback for speech sound errors in childhood apraxia. *Frontiers in Human Neuroscience, 10*(440), 440. https://doi.org/10.3389/fnhum.2016.00440
Preston, J. L., Leece, M. C., & Maas, E. (2017). Motor-based treatment with and without ultrasound feedback for residual speech-sound errors. *International Journal of Language & Communication Disorders, 52*(1), 80–94.
Preston, J. L., McCabe, P., Rivera-Campos, A., Whittle, J. L., Landry, E., & Maas, E. (2014). Ultrasound visual feedback treatment and practice variability for residual speech sound errors. *Journal of Speech, Language, and Hearing Research, 57*(6), 2102–2115. https://doi.org/10.1044/2014_JSLHR-S-14-0031
R Core Team. (2015). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/
Ruscello, D. M. (1995). Visual feedback in treatment of residual phonological disorders. *Journal of Communication Disorders, 28,* 279–302.
Ruscello, D. M., & Shelton, R. L. (1979). Planning and self-assessment in articulatory training. *Journal of Speech and Hearing Disorders, 44*(4), 504–512. https://doi.org/10.1044/jshd.4404.504
Rvachew, S. (1994). Speech perception training can facilitate sound production learning. *Journal of Speech and Hearing Research, 37,* 347–357.
Rvachew, S., Hodge, M., & Ohberg, A. (2005). Obtaining and interpreting maximum performance tasks from children: A tutorial. *Journal of Speech-Language Pathology and Audiology, 29*(4), 146–157.
Rvachew, S., & Jamieson, D. G. (1989). Perception of voiceless fricatives by children with a functional articulation disorder. *Journal of Speech and Hearing Disorders, 54*(2), 193–208.
Rvachew, S., Nowak, M., & Cloutier, G. (2004). Effect of phonemic perception training on the speech production and phonological awareness skills of children with expressive phonological delay. *American Journal of Speech-Language Pathology, 13*(3), 250–263.
Rvachew, S., Ohberg, A., Grawburg, M., & Heyding, J. (2003). Phonological awareness and phonemic perception in 4-year-old children with delayed expressive phonology skills. *American Journal of Speech-Language Pathology, 12*(4), 463–471.
Rvachew, S., Rafaat, S., & Martin, M. (1999). Stimulability, speech perception skills, and the treatment of phonological disorders. *American Journal of Speech-Language Pathology, 8*(1), 33–43.
Schmidt, A. M. (2007). Evaluating a new clinical palatometry system. *Advances in Speech-Language Pathology, 9*(1), 73–81.
Schmidt, R. A., & Lee, T. D. (2011). *Motor control and learning: A behavioral emphasis* (5th ed.). Champaign, IL: Human Kinetics.
Shriberg, L. D. (2009). Childhood speech sound disorders: From postbehaviorism to the postgenomic era. In R. Paul & P. Flipsen (Eds.), *Speech sound disorders in children*. San Diego, CA: Plural Publishing.
Shriberg, L. D., & Kwiatkowski, J. (1987). A retrospective study of spontaneous generalization in speech-delayed children. *Language, Speech, and Hearing Services in Schools, 18*(2), 144–157. https://doi.org/10.1044/0161-1461.1802.144
Shuster, L. I. (1998). The perception of correctly and incorrectly produced /r/. *Journal of Speech, Language, and Hearing Research, 41,* 941–950.
Shuster, L. I., Ruscello, D. M., & Smith, K. D. (1992). Evoking [r] using visual feedback. *American Journal of Speech-Language Pathology, 1*(3), 29–34.
Silverman, F. H., & Paulus, P. G. (1989). Peer reactions to teenagers who substitute /w/ for /r/. *Language, Speech, and Hearing Services in Schools, 20,* 219–221.
Skelton, S. L., & Hagopian, A. L. (2014). Using randomized variable practice in the treatment of childhood apraxia of speech. *American Journal of Speech-Language Pathology, 23*(4), 599–611. https://doi.org/10.1044/2014_AJSLP-12-0169
Thoonen, G., Maassen, B., Gabreels, F., & Schreuder, R. (1999). Validity of maximum performance tasks to diagnose motor speech disorders in children. *Clinical Linguistics & Phonetics, 13*(1), 1–23.
Thoonen, G., Maassen, B., Wit, J., Gabreëls, F., & Schreuder, R. (1996). The integrated use of maximum performance tasks in differential diagnostic evaluations among children with motor speech disorders. *Clinical Linguistics & Phonetics, 10*(4), 311–336.
Unicomb, R., Colyvas, K., Harrison, E., & Hewat, S. (2015). Assessment of reliable change using 95% credible intervals for the differences in proportions: A statistical analysis for case-study methodology. *Journal of Speech, Language, and Hearing Research, 58*(3), 728–739.
Van Riper, C. (1963). *Speech correction: Principles and methods* (4th ed.). Englewood Cliffs, NJ: Prentice Hall.
Wagner, R. K., Torgesen, J. K., Rashotte, C. A., & Pearson, N. R. (2013). *Comprehensive Test of Phonological Processing–Second Edition*. Austin, TX: Pro-Ed.
Wechsler, D. (2011). *Wechsler Abbreviated Scales of Intelligence–Second Edition*. San Antonio, TX: Pearson.
Wiig, E., Semel, E., & Secord, W. (2013). *Clinical Evaluation of Language Fundamentals–Fifth Edition*. Bloomington, MN: Pearson.

Appendix

Sample Schedule of Treatment Program

Monday a.m.	Tuesday a.m.	Wednesday a.m.	Thursday a.m.	Friday a.m.
2.5-hr testing Probes Overview of ultrasound	Session 2 15-min break Session 3	Session 6 15-min break Session 7	Session 9 15-min break Session 10	Session 13 15-min break Session 14
Monday p.m.	**Tuesday p.m.**	**Wednesday p.m.**	**Thursday p.m.**	**Friday p.m.**
Session 1	Session 4 15-min break Session 5	Probes Session 8	Session 11 15-min break Session 12	1-hr testing and wrap-up Probes