HHS Public Access
Author manuscript
J Commun Disord. Author manuscript; available in PMC 2018 April 14.

Published in final edited form as:
J Commun Disord. 2016 ; 64: 62–77. doi:10.1016/j.jcomdis.2016.10.003.

Acquisition, Retention, and Generalization of Rhotics with and without Ultrasound Visual Feedback

Greta M. Sjolie[1], **Megan C. Leece**[1], and **Jonathan L. Preston**[1]

[1]Department of Communication Sciences & Disorders, Syracuse University, 621 Skytop Rd, Suite 1200, Syracuse, NY 13244

Abstract

Purpose—The purpose was to provide a preliminary within-participant comparison of speech therapy with and without exposure to ultrasound visual feedback for postvocalic rhotics (/r/-colored vowels). Effects of the two treatments on acquisition, retention, and generalization were explored. It was hypothesized that treatment with ultrasound would facilitate acquisition but hinder retention and generalization.

Methods—A single subject randomized block design was replicated across four American English-speaking participants ages 7–9 years. Each participant was trained on postvocalic /r/. Each week for seven weeks, one session with ultrasound visual feedback and one session with no ultrasound were randomly ordered. A Training Probe and Generalization Probe were used to measure acquisition within each session as well as retention and generalization between two consecutive sessions. Graphical displays of the data, effect size calculation, and statistical results from a randomization test were used to analyze the results.

Results—Two participants showed essentially no evidence of acquisition, retention or generalization of rhotics (<5%). Of the two who showed evidence of acquisition, one participant showed a significant advantage and large effect size for ultrasound sessions over no ultrasound sessions in acquisition of rhotics. However, no participants showed differences between treatment conditions in generalization or retention of rhotics.

Conclusion—For some children, acquisition may be facilitated by ultrasound visual feedback. Ultrasound visual feedback neither inhibited nor facilitated retention or generalization of rhotics. As a whole, the 14 treatment sessions (7 with ultrasound and 7 without) were effective for 2 of the 4 participants when comparing pre/post generalization scores. Future studies should evaluate the effectiveness of ultrasound visual feedback given a larger dose and differing age groups.

Keywords

ultrasound; visual feedback; rhotics; speech sound disorders; speech therapy

---

Corresponding Author: Dr. Jonathan Preston, Department of Communication Sciences & Disorders, Syracuse University, 621 Skytop Rd, Suite 1200, Syracuse, NY 13244, jopresto@syr.edu/ 315-443-3143.

1. Introduction

The /ɹ/ phoneme is among the most common sounds in error in American English-speaking preschool and school-age children with speech sound disorders (e.g., Shriberg, 2009). Typically, American English /ɹ/ is one of the last phonemes to be acquired (Smit, Hand, Freilinger, Bernthal, & Bird, 1990) possibly due to its phonetic complexity (Boyce, 2015). It is important for SLPs to rely on evidence-based techniques when making clinical decisions to help children resolve speech errors. Moreover, testing treatment outcomes may aid in the identification of effective and efficient treatments that enable more rapid progress, as unresolved speech errors have been reported to contribute to academic and socioemotional difficulties (e.g., Lewis, Freebairn, & Taylor, 2000; Shriberg & Kwiatkowski, 1988). The present study focuses on the effects of treatment with and without ultrasound visual feedback of the tongue on acquisition, retention, and generalization for children ages 7–9 whose /ɹ/ distortions have not yet resolved.

1.1 Lingual Components Necessary for /ɹ/

American English rhotics are phonetically complex and may appear in several positional contexts (onset, nucleus, or coda) as vocalic or consonantal segments. Regardless of the syllable position, the following lingual components are typically necessary for the proper production of /ɹ/: a pharyngeal constriction (tongue root retraction to create posterior narrowing of the vocal tract), an oral constriction (achieved by tongue tip retroflex or bunching of the tongue body), tongue midline grooving, and (in most speakers) elevation of the lateral margins of the tongue (Bernhardt et al., 2008; Bernhardt, Gick, Bacsfalvi, & Adler-Bock, 2005). Anterior tongue shapes for /ɹ/ are typically divided into two general categories: retroflex and bunched. For a retroflex /ɹ/, the tongue tip is raised and may be curled back slightly; for a bunched /ɹ/, the tip of the tongue lowers, while the dorsum rises toward the hard palate (Klein, McAllister-Byun, Davidson, & Grigos, 2013).

Although these components are typically observed for correct /ɹ/ production, individual variation exists; for example, midsagittal MRI images reveal a variety of tongue shapes with differing amounts of pharyngeal constriction and midline grooving to produce a perceptually correct /ɹ/ (e.g., Boyce, 2015). Generally, an accurate /ɹ/ cannot be produced with only one component (Gick et al., 2007). Recently, research has revealed that distorted /ɹ/ productions are associated with a highly curved, posteriorly located tongue shape which likely results in loss of pharyngeal constriction (Adler-Bock, Bernhardt, Gick, & Bacsfalvi, 2007; Boyce, 2015).

1.2 Traditional Therapy for /ɹ/

Some common methods for /ɹ/ therapy include imitation, contextual facilitation, phonetic placement and shaping, motokinesthetic training, and perceptual awareness training (Adler-Bock et al., 2007; Secord, Boyce, Donohue, Fox, & Shine, 2007). For example, Shriberg (1975) described a multistep approach for shaping /l/ to /ɝ/. Later, Shriberg (1980) developed a procedure to eliminate incorrect articulatory movements for /ɹ/ using a bite stick. Some children improved their articulation of /ɹ/; however, the treatment did not always result in correct productions.

Verbal instructions encouraging children to curl the tongue tip up or bunch the tongue may be used to elicit /ɹ/, although these cues may be abstract and difficult to describe to children. Lateral contact of the sides of the tongue against the molars is the only tactile feedback available for /ɹ/, and this offers no specific information about where to place the front of the tongue in order to achieve the desired constrictions.

While these approaches to /ɹ/ therapy are still used today, they do not always yield a perceptually correct /ɹ/. Additionally, increasing access to new technologies warrants consideration for enhancing the effectiveness of /ɹ/ therapy as clinical observations reveal that numerous children remain in /ɹ/ therapy for years, or are dismissed from therapy on the grounds that no improvements have been made. Therefore, when traditional approaches to /ɹ/ therapy have not worked for a child, alternative approaches such as incorporating technology into therapy should be explored (Adler-Bock et al., 2007; Ruscello, 1995).

1.3 Visual feedback Therapy for /ɹ/

Visual feedback involves the presentation of information that allows individuals to gain greater awareness and control over physiological functions (Volin, 1998). In regards to speech, visual feedback enables aspects of speech that are usually difficult to perceive (i.e., tongue movement) to be brought under conscious control using information from real-time instrumental feedback (McAllister-Byun, Hitchcock, & Swartz, 2014; Ruscello, 1995; Volin, 1998). One common application of biofeedback in traditional speech therapy is the use of a mirror to provide visual feedback on the movement of the lips and jaw during speech articulation. However, the lingual gestures of /ɹ/ cannot be easily seen with a mirror due to the presence of lip rounding and a relatively small oral cavity opening; thus, clinicians are forced to rely on acoustic information and clinical experience to cue the desired tongue shape (Secord et al., 2007)

1.3.1 Ultrasound Imaging of the Tongue as Visual Feedback—Ultrasound (US) in speech therapy has been shown to be safe and to have no detrimental side effects (Epstein, 2005; Preston et al., 2014). The transducer is placed beneath the chin and allows for an image of the upper contour of the tongue to be displayed. US reveals movements in real time (or captured at certain points in time) that can help the clinician select appropriate cues to facilitate a correct production of /ɹ/ and allow children to visualize the exact movement of the tongue. In a mid-sagittal view, the anterior lingual constriction (produced by the tongue tip or blade) and the posterior constriction (tongue root retraction toward the pharyngeal wall) are visible. Lateral elevation and depth of the midline groove are also visible with US using a coronal view (Bernhardt et al., 2005).

US visual feedback has been shown to facilitate more accurate /ɹ/ production in a variety of clients (e.g., Adler-Bock et al., 2007; Bacsfalvi et al., 2007; Bacsfalvi & Bernhardt, 2011; Bernhardt, Gick, Bacsfalvi, & Ashdown, 2003). Bernhardt et al. (2008) demonstrated that US can have positive benefits with less than three hours of actual practice in establishing sounds in speech therapy when clients possess motivation and have less pervasive residual speech impairments. Preston et al. (2014) studied eight participants with residual speech sound errors and reported that most participants increased accuracy for at least one treated sound during a treatment program including US visual feedback. McAllister-Byun et al. (2014) also demonstrated that an intervention including US visual feedback can be effective for children with persistent /ɹ/ errors. The above studies support the use of US in speech therapy; however, no studies have directly compared therapy with and without US in order to observe the benefits US may provide.

1.4 Principles of Motor Learning

As Shriberg (1980) noted, children with /ɹ/ distortions may have learned incorrect articulatory patterns that must be remediated. Therefore, approaching therapy from a motor learning framework may facilitate learning of the movements required for rhotics (e.g., Preston et al., 2014). Motor learning principles have recently been applied to treatment for children's /ɹ/ errors (Hitchcock & McAllister-Byun, 2015; Preston et al., 2014). It is unknown if the same principles of motor learning apply equally to non-speech motor tasks (e.g., limb movements) and speech motor tasks. However, the following hypothesis should be taken into consideration as evidence of effectiveness continues to be collected: principles of motor learning that apply to non-speech motor control may also apply to speech motor control (Maas et al., 2008). Thus, the following factors that have been shown to facilitate motor skill acquisition and learning in general are often considered in studies aimed at developing effective interventions for speech sound errors: structure of practice, stimulus selection, and nature of feedback (e.g., Hitchcock & McAllister-Byun, 2015; Maas et al., 2008; Preston et al., 2014; Skelton & Hagopian, 2014).

When principles of motor learning are applied to speech tasks, an important distinction can be made between acquisition of a skill, retention of an acquired skill, and generalization of an acquired skill. As defined by schema-based motor learning theory (Maas et al., 2008; Schmidt & Lee, 2011), *acquisition* refers to successful attempts during practice (i.e., within session performance during training). In speech therapy, acquisition occurs as trained speech sounds are repeatedly practiced and shaped into new, more accurate movements. Thus, evidence of acquisition may be reflected in short-term changes in accuracy of a speech sound from the beginning to the end of a session. Although some degree of skill acquisition is necessary before learning can occur, acquisition does not imply motor *learning*. Motor learning must be observed through retention and generalization (Maas et al., 2008; Schmidt & Lee, 2011).

Retention refers to performance levels after practice is completed. In speech therapy, retention can be observed by determining if trained items containing a target sound practiced during a session are correctly produced at a later time. *Generalization* refers to how practice on one movement affects similar, but untrained, movements. In speech therapy, generalization may be observed by determining the accuracy of untrained items containing the target sound that are correctly produced after a duration of time. Therefore, speech therapy aims to ultimately maximize learning (retention and generalization), not acquisition, during treatment sessions (Maas et al., 2008) as effective interventions for speech sound errors does not solely measure success on trained tasks, but expects to see generalization to similar, untrained tasks. To date, no study has examined the effects of acquisition and learning in therapy by comparing sessions that included the frequent use of US compared to no US.

1.4.1 Feedback—

Feedback is vital to establishing a new motor plan and two *types of feedback* are typically distinguished: knowledge of performance (KP) and knowledge of results (KR). KP feedback refers to feedback related to how the movement was produced, while KR feedback refers to feedback related to the results produced in terms of the goal (e.g., correct or incorrect). Traditional approaches to /ɹ/ therapy deliver KP feedback via verbal instruction (e.g. “Raise the front of your tongue a little more”). In addition to verbal KP feedback by the clinician, US provides KP feedback visually by allowing children to observe the shape and position of their tongues; therefore, two distinct modalities of KP feedback can be provided to help facilitate motor performance (Maas et al., 2008 & Preston et al., 2014).

KP feedback has been found to facilitate the rate of acquisition and overall performance level when a non-speech task is unfamiliar (Newell, Carlton, & Antoniou, 1990). However, Hodges and Franks (2001) demonstrated that an increase in KP feedback during non-speech tasks may impede motor learning, possibly because learners can become over-reliant on KP feedback. The effects of KP feedback using US as visual feedback on acquisition and generalization of speech sounds have not yet been established in previous research.

1.4.2 Practice—

The *complexity of stimuli* practiced in speech therapy is important to consider as children may learn best when they are appropriately challenged at a level commensurate with their ability (Guadagnoli & Lee, 2004; Hitchcook & McAllister Byun, 2015). Selected stimuli should increase in complexity to a level consistent with a child’s motor performance to increase the task difficulty.

An additional consideration in the structure of speech treatment is *practice schedule* (Maas et al., 2008). Blocked practice refers to a practice schedule wherein the client practices a group of the same target movements (e.g., AAA, BBB, CCC) and random practice refers to a practice schedule wherein different movements are successively practiced (e.g., ABC, CBA, and BCA). Based on predictions from schema-based motor learning theory, blocked practice may enhance acquisition while random practice may enhance motor learning for intact and impaired speech motor systems (Maas et al., 2008; Skelton & Hagopian, 2014); thus, clients may benefit from starting with a blocked practice schedule and then transitioning to a random practice schedule.

1.5 Present Study

Although a number of single case experimental studies have been conducted exploring the efficacy of US visual feedback, studies have not yet directly compared the use of treatment with and without US. Learning (i.e., retention and generalization) is the primary goal of speech therapy, but acquisition is a necessary prerequisite to learning. Thus, the present study examined how US as visual feedback treatment compared to no US treatment in facilitating performance within a session to trained targets, retention between consecutive sessions to trained targets, and generalization between sessions to untrained targets. A motor learning framework is used to guide the treatment structure.

As visual KP feedback may facilitate acquisition (Newell et al., 1990) and impede learning (Hodges & Franks, 2001) during non-speech tasks, it is possible that visual feedback from US may enhance acquisition during speech. Therefore, it was hypothesized that, compared to treatment without US visual feedback,

1. treatment with US visual feedback will facilitate acquisition of rhotics
2. treatment with US visual feedback will impede retention of rhotics,
3. treatment with US visual feedback will impede generalization of rhotics.

2 Methods

2.1 Participant Characteristics

Four participants were recruited through referrals from local SLPs as part of a larger study on ultrasound treatment. Individuals with the following diagnosed disabilities were excluded from this study (as determined by a phone screening): developmental disabilities such as autism or cerebral palsy, hearing loss, known or likely brain injury; underlying structural or functional abnormalities such as cleft palate; known genetic syndromes such as Down syndrome. Individuals diagnosed with or suspected of having childhood apraxia of speech (CAS) or dysarthria during qualifying assessments were also excluded.

To be included in the study, participants had to show evidence of a speech sound disorder based on the sounds-in-words subtest of the Goldman-Fristoe Test of Articulation-2 (GFTA-2; Goldman & Fristoe, 2000). Participants also had to score below 30% on rhotic diphthongs on probe lists (described below) in at least two of the following four contexts: /ɑɹ ɪɹ ɔɹ ɛɹ/.

Additional inclusionary criteria included the following: between the ages of 7;0 to 9;11; monolingual American English-speakers (speaking a dialect with postvocalic /ɹ/); hearing within normal limits as evidenced by passing a screening bilaterally at 20 dB at 1, 2, and 4 kHz; visual acuity in one eye of at least 20/40 based on the Snellen Eye Chart; and scores no lower than 1 SD below the mean on the Matrix Reasoning Index score of the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 2011). Language scores had to be broadly within normal limits based on standard scores above 80 on the Peabody Picture Vocabulary Test-4 (PPVT-4; Dunn & Dunn, 2007) and scaled scores greater than 7 on the Recalling Sentences and Formulated Sentences subtests on the Clinical Evaluation of Language Fundamentals-5 (CELF-5; Wiig, Semel, & Secord, 2013). Descriptive data can be found in Table 1.

2.1.1 Descriptive Measures—During pre-treatment baseline sessions, additional testing took place for descriptive purposes. To characterize phonological processing skills, the Elision and Blending Words subtests of The Comprehensive Test of Phonological Processing-2 (*CTOPP-2*; Wagner, Torgesen, & Rashotte, 2013) were administered. The Speech Assessment and Interactive Learning System (SAILS) (Rvachew, 1994) was also administered to assess participants' ability to perceptually judge productions of phonemes as correct or incorrect; twenty tokens each of /ɹ s f ʃ θ/ were administered. To evaluate phonological working memory, a Nonword Repetition Task (Dollaghan & Campbell, 1998) was administered. A Percent Phonemes Correct (PPC) out of 96 total phonemes was computed.

A Maximum Performance Task (cf. Thoonen, Maassen, Wit, Gabreels, & Schreuder 1996, 1999; Rvachew, Hodge, & Ohberg, 2005) was administered in order to evaluate motor speech function in sustained phonemes (e.g., /ɑː/), repeated monosyllables (e.g., /pʌpʌpʌ/) and multisyllables (e.g., /pʌtʌkʌ/). A Multisyllabic Word Repetition Task (Preston & Edwards, 2007) was administered to assess sound sequencing, and syllable-to-syllable transitioning of sounds. Participants' performances on the above speech assessments were used to rule out significant motor speech impairment.

Finally, a stimulability probe for /ɹ/ (adapted from Miccio, 2002) was administered to characterize participants' pre-treatment rhotic accuracy in 11 imitated syllables, each repeated three times. Each attempt was scored as correct or incorrect based on the rhotic. This was completed with the ultrasound to allow the clinician to determine pre-treatment tongue shape and to consider the type of tongue placement cues that might be necessary during therapy.

2.2 Data Collection

Two probe lists were administered during assessment sessions prior to the initiation of therapy and also at the start of each therapy session: a Training Probe (comprised of 10 items that were treated per condition, or 20 total) and a Generalization Probe (comprised of 25 untreated items). The Training Probe and the Generalization Probe were combined into one randomized master list containing 45 words, and this was administered imitatively due to the young ages of the participants. Imitation was used as some training probes reached a 6–7 syllable length and included multisyllabic words above a 7 year old reading level. The combined probe list was administered at the beginning of each session in a different order. At the end of the session, only the Training Probe was administered to evaluate acquisition of treated items.

Probes assessing accuracy of other phonemes in error (in this instance, /s/) were administered to serve as untreated controls. Probes were audio recorded in Praat (Boersma & Weenink, 2014) for later scoring.

2.2.1 Acquisition (Training Probes)—Administration of the same Training Probe at the beginning and the end of each session allowed for acquisition to be quantified; change in accuracy on these probes each session served as the dependent variable for hypothesis 1. Two postvocalic rhotic contexts (/Vɹ/combinations), which included the participant's least accurate context selected from /ɑɹ, ɛɹ, ɪɹ, and ɔɹ/ were trained for each participant. Each postvocalic rhotic context included 10 initial training words: five which were randomly selected to be trained during US sessions and five which were trained during no ultrasound (NoUS) sessions, where any given word was only trained under one of the conditions. Therefore, only 10 words were trained during each session and acquisition was quantified by changes in accuracy on the Training Probe items from the beginning to the end of the session (based on the average of the four raters from recordings). The use of different words for the two treatment conditions was designed to prevent contamination between items trained in the two treatment conditions (see Appendix A for a sample Training Probe List).

The number of Training Probe items incorrectly produced at the beginning of the session provided the number of probes available for improvement out of 10 (i.e., 10 Training Probe items for that session). A percentage was calculated by subtracting the number of correctly produced Training Probe items perceptually judged by the treating clinician out of 10 at the beginning of the session from the number of correctly produced Training Probe items out of 10 at the end of the session.

All items trained within a session were taken from the Training Probe list. Each Training Probe item began at the monosyllabic level (see Appendix B for sample Training Probe Chains); if a participant correctly produced the monosyllabic Training Probe item at the beginning of the session, a multisyllabic word replaced the monosyllabic Training Probe. If the multisyllabic word was also correctly produced, then a varied longer phrase (i.e., a phrase or short sentence containing six to seven syllables) including the multisyllabic word replaced the multisyllabic word in the Training Probe list. This helped to ensure that appropriately complex items were trained. Additionally, to avoid training a phrase that had already been correctly produced during administration of the Training Probe at the beginning of the treatment session, a word bank of additional words and phrases was available so additional items could be probed and trained (Sjolie, 2015). Ten items were always trained to keep the number of training items consistent across sessions, and the percent of items acquired in the session was the dependent variable for measuring acquisition.

2.2.2 Retention (Training Probes)—Administration of the same Training Probe items at the beginning of the next consecutive session allowed for retention to be quantified; accuracy on these probes served as the dependent variable for hypothesis 2. Retention refers to performance levels after practice is completed; thus, the assumption for this study was that retention was attributed to the prior treatment session. Therefore, the percentage of the 10 Training Probe items trained in the previous session that were correctly produced by participants at the beginning of the next session was calculated to reflect retention.

2.2.3 Generalization—Accuracy on the 25 Generalization Probe items administered at the beginning of each session served as the dependent variable for Hypothesis 3. The Generalization Probe consisted of 25 items including monosyllabic words, multisyllabic words, phrases, and short sentences; each contained one scored rhotic which included untrained words and untrained sound contexts (/ɚ/, /Vɹ/, /ɹV/, /Cɹ/). Generalization was observed by comparing the percentage of correctly produced Generalization Probe items between two successive sessions. This difference accounted for the amount of generalization that occurred from the beginning of one session to the beginning of the next session. The assumption was that gains in generalization were attributable to the prior treatment session.

2.3 Study Design

Each week involved two different treatment sessions for the participant: one ultrasound (US) and one no ultrasound (NoUS), which were randomly ordered each week. This single subject randomized block design was replicated across four subjects. A randomized block design is similar to an alternating treatment design except that within each block (each week) the treatment order is randomly determined instead of remaining the same (Kratochwill & Levin, 2014). Randomization allows for statistical significance testing (Rvachew, 1988). This type of assignment method guaranteed that conditions of the same type could not appear in more than two consecutive sessions (Kratochwill & Levin, 2014). Participants attended two sessions a week for a total of 14 sessions over a seven to eight week period. If participants received speech therapy from another SLP while participating in this study, the SLP was contacted and agreed to not treat /ɹ/ during the study.

Participants were randomly assigned 3–5 baseline probes which were collected approximately twice per week during sessions in which the other speech and language tasks were administered. After treatment was complete, participants returned for two follow-up sessions. One session took place approximately one week after treatment ended to collect post-treatment data on the final Training Probe, the Generalization Probe, and the Untreated Sound Probe. Additionally, one month after treatment ended, participants returned to evaluate long-term retention and generalization.

2.4 Treatment Target Selection

The baseline Training Probe list contained five words for each of the /Vɹ/ contexts occurring in any word position. Two rhotic contexts to be trained during therapy (e.g., /ɑɹ/ and /ɪɹ/) were selected based on contexts the participant produced with 30% accuracy or lower when averaged over all baseline probes.

2.5 Treatment Overview

All four participants were treated by the same speech-language pathology graduate student under supervision of a certified speech-language pathologist.

Treatment sessions were adapted from those outlined by Preston et al. (2014). US treatment and NoUS treatment always began with administration of the Generalization and Training Probes. Each session included three 13 min time periods of therapy (A, B, C). Between each Time Period a 2–3 min game was played to provide a break. A timer was used to ensure adherence to the allotted time. US treatment and NoUS treatment were identical in target selection methods, treatment duration, practice schedule, and feedback frequency. The difference between conditions was the use of the ultrasound for 26 min in US treatment sessions (13 min in Time Periods A and C) which increased the specificity of KP feedback allowing the clinician to provide specific tongue placement cues when using verbal KP feedback and allowing participants access to visual feedback. Table 2 outlines the US treatment and NoUS treatment session structure.

2.5.1 Condition differences—During Time Periods A and C of US treatment sessions, practice attempts of vocalic targets occurred with both verbal and visual feedback. A Seemore PI 7.5 MHz or Echo Blaster 128 with a PV6.5/10/128Z transducer was used for tongue imaging. Both sagittal and coronal views were used at the discretion of the treating clinician. Similar to prior studies (e.g., Preston et al., 2014, 2015), cues were provided based on the child's production in an attempt to achieve an oral constriction with the anterior tongue (e.g., "lift this part of your tongue up"), achieve a pharyngeal constriction (e.g., "move this part of your tongue back"), achieve lateral elevation of the sides of the tongue (e.g., "lift the sides of your tongue up"), and inhibit incorrect movement (e.g., "keep the body of your tongue lower").

During NoUS sessions, and during Time Period B of US treatment sessions where no US was provided, productions of rhotics continued but without the visual display. Drill-like activities were used and verbal tongue placement cues were provided without reference to any visual display (e.g., "touch the sides of your tongue to the inside of your top teeth"). During NoUS sessions, all three Time Periods (A, B, and C of NoUS) involved treatment without the ultrasound.

2.5.2 Elicitation (Pre-practice)—Regardless of whether the session was US or NoUS, each session began in the elicitation (pre-practice) phase during Time Period A. The participant had to accurately produce each of the two target rhotic diphthongs six times (judged by the clinician) before structured practice could begin. During elicitation, any amount and type of cueing or feedback was allowed (whereas during structured practice, the amount, type of feedback, and practice was predetermined). Verbal models, shaping techniques (e.g., shaping [ɹ] from [l] or [ɑ]), phonetic placement cues, drawings, and contextual facilitation were used to encourage an acceptable /ɹ/. If a participant did not produce six accurately judged target sounds of each context by the clinician during Time Period A, then the elicitation phase continued during the next Time Periods until this criterion was met. It was possible for a participant to remain in the elicitation phase for the entire duration of a therapy session. However, regardless of whether a participant remained in the elicitation phase, the specifications regarding each Time Period as an US or NoUS session shown in Table 2 were upheld (e.g., time duration and use of US).[1]

2.5.2 Structured Practice—Each target sound was trained using chaining to increase complexity: building from simple to complex stimuli. The criterion of 5 of 6 correct productions was established for participants to progress from syllables to monosyllabic words, multisyllabic words, short phrases, and varied longer phrases. Once participants correctly produced all levels of a chain, new chains were added (which may have included multiple rhotics in the phrase).

Both US and NoUS treatment conditions began with blocked practice and ended with random practice. Blocks of six trials were used to practice the target sound. When participants reached the varied long phrase level, the position in the sentence of the target sound was varied.

Verbal KP feedback and verbal KR feedback were systematically provided during Structured Practice. High frequency KP and KR feedback was used when establishing the rhotic vowel target at the syllable level. However, as participants progressed to more complex targets,

[1]Two participants (1003, 1008) met the criterion of 12 correct attempts to pass from elicitation to structured practice during all 14 sessions. Participant 1004 met the criterion first during session 7 (a US session), and met the criterion only during 7 out of 14 sessions. Participant 1010 met the criterion only during sessions 11 and 14, both US sessions.

feedback was faded and was reduced to primarily KR feedback (see Table 3). Verbal KP feedback and verbal KR feedback were randomly assigned to each trial (e.g., KR and KP provided for first trial vs. only KR provided for first trial) within each block of six trials.

Random practice was provided during the last six minutes of time period C. Thus, Time Period C in all sessions included part C-I lasting 7 minutes, which included typical chaining procedures and C-II lasting 6 minutes which included random practice (see Table 2). However, if the participant did not move past the elicitation phase (described in 2.5.1), elicitation continued during C-II. Random practice occurred on the highest linguistic level reached in each chain during training (i.e., monosyllabic, multisyllabic, short phrase, or varied long phrase) and involved randomly presenting all targets at that level.

2.6 Treatment Reliability and Fidelity

Training Probe items and Generalization Probe items were scored by four listeners (the treating clinician and three additional listeners who were unaware of treatment conditions) via audio recordings. All listeners were provided with written guidelines and a training on the criteria to score a rhotic as correct or incorrect. In general, listeners were instructed to score "marginally correct" productions as incorrect. Lists were randomized, and scoring was then blocked, with listeners scoring each item from a probe as correct or incorrect. The percentage of probes scored as correct by each of the four listeners were averaged as a whole and used as the percent of probes correct for the final data analysis.

Scores were compared between listeners to determine the degree of agreement. The absolute value between all four listeners' scores was computed for acquisition, retention, and generalization for each participant (including baseline and follow-up probes). The average pairwise absolute difference between all four pairs of listeners was 7.2% (SD 7.8).

Treatment fidelity checks were performed by a research assistant who reviewed two randomly selected sessions per participant (one in each treatment condition) via audio or video recordings. Adherence to verbal KR and feedback schedule, calculated per trial for all eight sessions was 96.3%.

Additionally, to ensure the treating clinician provided appropriate feedback on the accuracy of /ɹ/, reliability judgments were completed on every trial where explicit KR feedback was provided by the treating clinician. Agreement on the appropriateness of KR feedback between the treating clinician and the research assistant was 91.1% (SD 3.0).

Moreover, for sessions selected for the fidelity check the number of trials elicited from a participant during elicitation and structured practice within a session was computed. The average number of rhotic trials per session was 215 (SD 51).

2.7 Data Analysis

The independent variable was the treatment condition which had two levels: US and NoUS. There were three dependent variables: probe scores reflecting acquisition, retention, and generalization. To evaluate outcomes, graphical displays of the data were paired with descriptive statistics, effect size calculation (Cohen's d, which is the mean difference between conditions divided by the pooled standard deviation), and statistical results from a randomization test (Kratochwill & Levin, 2014). The test statistic was the mean difference between US sessions and NoUS sessions which was computed for each week of therapy. A randomization test using all possible permutations of data was used to obtain p-values using the Single Case Randomization Test in the statistical software R (R Core Team, 2015). The p-value corresponding to the observed mean difference is obtained by ranking the obtained value in the distribution of all assignment possibilities (Bulte & Onghena, 2008). The p-value represents the proportion of possible test statistics in the distribution that are above the observed test statistic.[2] Each p-value is then compared with a critical value for significance testing where the proposed study had a critical value of alpha=.05. Given the exploratory nature of this single case study and the limited possible p-values with a randomization test, an a priori decision was made that multiple corrections for the alpha level were not used.

3 Results

3.1 Acquisition

Figure 1 shows participants' percent change in acquisition each session (change in accuracy on the Training probe administered at the beginning and end of the session). Figure 1 suggests that 1 of 4 participants showed an advantage for US sessions over NoUS sessions during acquisition. Stable baselines were observed for all participants. Participant 1003 showed a generally consistent advantage for US sessions over NoUS sessions. Participant 1008 showed signs of acquisition, but no consistent advantage for either US sessions or NoUS sessions. Acquisition scores for participants 1004 and 1010 were at or near 0% in all sessions, suggesting no advantage for either treatment condition.

Table 4 presents descriptive statistics, effect sizes, and p-values associated with the acquisition data. Consistent with the graphical trend, participant 1003 showed a significant advantage of US sessions over NoUS sessions in acquisition scores ($p=.039$, $d=.78$); however, the remaining three subjects did not show a significant advantage for either treatment.

3.2 Retention

Figure 2 shows participants' percent of retention for probes following each treatment condition (i.e., percent of items that were trained in the previous session that were correctly produced at the start of the subsequent session). Figure 2 revealed a negligible difference between US sessions and NoUS sessions. Participants 1004 and 1010 showed retention scores primarily at 0%: there was not an advantage for either treatment condition.

Descriptive statistics, effect sizes, and p-values are shown in Table 5. None of the participants showed a statistically significant advantage for one treatment over the other in retention scores.

[2]Although the hypotheses are directional, to be comprehensive two p-values per condition are reported: one p-value testing the assumption that the advantage was for the US session and the other testing the assumption that the advantage was for the NoUS session.

3.3 Generalization

Figure 3 displays the percent of Generalization Probe items produced correctly following each treatment condition. Baselines were stable (within a range of <12%) for all participants. Figure 3 revealed no apparent advantage for NoUS or US conditions. Participant 1003 and 1008 demonstrated an upward trend in generalization but a minimal difference between conditions. Participants 1004 and 1010 made negligible gains in generalization. Additionally, descriptive statistics, effect sizes, and p-values are shown below in Figure 3 and in Table 6.

3.4 Untreated Sounds

Participant 1004 and 1010 both had word-final /s/ distortions that were monitored as untreated controls. Figure 4 displays the percent correct of /s/ words over the course of the study. No upward trend is evident, suggesting no apparent maturation effects during the duration of the study.

4 Discussion

This study compared effects of US treatment and NoUS treatment for 7–9 year olds with rhotic distortions. It was predicted that the use of US as KP visual feedback would facilitate acquisition within a session; however, the use of US as KP visual feedback would hinder motor learning in the form of retention and generalization.

4.1 Acquisition

The increase in accuracy from beginning to the end of each session served as the dependent variable to address hypothesis 1. Two of four participants showed evidence of acquisition on the probes whereas two did not show signs of acquisition. Among the two who showed signs of acquisition, one participant (1003) showed a statistically significant advantage for US treatment over NoUS treatment (d=.78) whereas the other participant (1008) showed no apparent advantage for either treatment (d=.13). Neither participant 1004 nor 1010 made notable progress in acquisition throughout the 14 sessions.

Consequently, hypothesis 1 was supported by data from only one of the four participants. Therefore, these findings provided some support, but not strong support, for the theoretical assertion that visual KP feedback facilitates speech sound acquisition. Visual KP feedback was not observed to be detrimental to acquisition for any participant. In non-speech tasks that are not well established, KP feedback may facilitate the rate of acquisition and overall performance levels (Newell et al. 1990); additionally, the present study suggested visual KP feedback may be useful for some individuals in facilitating acquisition of speech sounds.

4.2 Motor Learning

Motor learning was evaluated through probes assessing retention of trained rhotics and generalization of untrained rhotics. Neither participant 1004 or 1010 made notable progress in retention or generalization throughout the 14 treatment sessions, which is expected given their limited acquisition of the rhotic targets. Therefore, only data from participants 1003 and 1008 could be used to evaluate differences in motor learning between US and NoUS conditions.

Regarding hypothesis 2, no participants showed statistically significant differences in retention between the two treatments. When the effect sizes were examined, participant 1003 showed a slight advantage for US sessions over NoUS sessions (d=.48) whereas, participant 1008 showed a small advantage (d= −.30) for NoUS sessions over US sessions. However, there was no visually apparent distinction between the two conditions.

It was notable that participant 1003, who showed a strong advantage for US treatment during acquisition, showed the strongest (but not statistically significant) advantage for US treatment in retention. Although an increase in KP feedback through the use of visual feedback was predicted to impede motor learning (Hodges & Franks, 2001), for participant 1003, visual KP feedback was associated with moderately increased retention between sessions. Thus, increased KP feedback through additional use of ultrasound did not impede motor learning. Moreover, the data suggest that an increased rate of acquisition due to US treatment may also facilitate (and/or may not hinder) retention. Additional studies with larger sample sizes are needed to further examine this claim.

Regarding hypothesis 3, generalization was similar following both treatment conditions for all participants. Effect sizes between conditions were small, with mean differences between conditions of about 1%. This observation contradicts the suggestion that an increase in KP feedback may impede motor learning (Hodges & Franks, 2001). It is possible that, for motor learning to suffer as a result of KP, the additional visual feedback provided must be of sufficiently high frequency to encourage children to rely exclusively on feedback. However, the US sessions implemented here included periods of practice with and without visual feedback (i.e., there was not 100% visual feedback), which might have mitigated some of the potentially negative effects of KP visual feedback on motor learning.

4.3 Additional Considerations

A single subject randomized block design replicated across four subjects provided multiple opportunities for the intervention to demonstrate an effect. The treatment dose (i.e. 14 sessions) was similar to other studies (e.g., Adler-Bock et al., 2007; McAllister Byun & Hitchcock, 2014; Preston et al., 2014) which have shown both treatment responders and non-responders. The dose may not have been enough for all participants to demonstrate an effect for the treatment. Both participants 1004 and 1010 demonstrated establishment of the trained sound in isolation within the last five treatment sessions. Thus, had the treatment duration been longer than 14 sessions, these participants may have revealed additional information concerning the rate of acquisition or learning within US sessions and NoUS sessions. Further research evaluating the effectiveness of different dosages of US visual feedback therapy is needed.

It is important to note that all participants established the target sounds in isolation during therapy. Even participants 1004 and 1010, who were not stimulable for any rhotic before therapy, showed slight improvements in stimulability and were successfully capable of producing rhotics first with the US and then during NoUS treatment sessions. Comparisons of participants' scores on the pre-therapy stimulability probe, and scores on the 1 week follow-up stimulability probe revealed that all participants' scores increased (stimulability scores increased from 39 to 76% for 1003, 0 to 9% for 1004, 55 to 61% for 1008, and 0 to 18% for 1010) Thus, the therapy package that contained US and NoUS treatment was successful in enhancing stimulability for participants. US has been found to be particularly useful at the beginning of therapy to help establish the sound in error by allowing participants to view their tongue configuration during a correct rhotic production. Further research evaluating the effectiveness of when an US visual feedback component is introduced in therapy is needed. Follow up research may consider examining the effectiveness of US as visual feedback versus NoUS on stimulability, and it will be useful to determine other options for individuals who fail to achieve stimulability with US (e.g., acoustic visual feedback, electropalatography, perceptual training, etc.).

Individual response to treatment may be due to several factors. For example, one might hypothesize that speech perception skills may influence treatment response (e.g., Wolfe, Presley & Masaris, 2003), but there were not substantial differences in speech perception or phonological processing skills of the treatment responders (see Table 1). However, the two participants who demonstrated a response to the treatment as a whole (1003, 1008) were older than the two who did not (1004, 1010). A more mature speech learning system may be needed to take advantage of the visual feedback. One speculated reason for this is that the older participants were better able to interpret the visual feedback the US provided and incorporate the feedback into their own rhotic production. Both participants 1003 and 1008 (who improved in their generalization scores) were stimulable for rhotics (e.g., initial syllable position) prior to treatment, scoring above 30% on the stimulability probe in some contexts. In contrast, prior to therapy, participants 1004 and 1010 scored 0% on the stimulability probe. Consequently, the acquisition, retention, and generalization rates for participants 1004 and 1010's sessions were consistently below 10%. Thus, age, speech motor maturation, and/or stimulability may be important factors when examining the success of acquisition and generalization within US and NoUS rhotic speech therapy within a 14-session program; children who are stimulable for a target sound may experience a higher success rate during acquisition.

Additionally, the participants who only had one sound in error experienced a higher success rate during the treatment program than did the two participants who had another sound in error. This could again be related to severity of participants' speech sound disorders or maturity of their speech systems. Further research evaluating the effectiveness of US visual feedback therapy on differing age groups and severity levels is needed, and obtaining information about visual-motor skills may also help to understand individual differences in treatment response.

4.4 Caveats and Limitations

The randomized within block single-subject design was used to avoid cumulative effects on one treatment type occurring three or more times together, and to control for rising trends in the data. However, this design is not without limitation as it cannot control for cumulative effects that may occur from effective treatment as a whole. Thus, the assumption was that generalization was attributed to the prior treatment session, though it is possible that the combined effects of all prior treatment could have influenced generalization. Additionally, given the nature of the design in which participants received both treatments, no conclusions could be drawn concerning the effectiveness of only US treatment or only NoUS treatment over the length of the study. Randomized group designs would be required to overcome these limitations.

It should also be noted that the comparison of interest in this study was between sessions with and without visual KP. Both treatment conditions included verbal KP, as is common in most motor-based treatment for speech sound disorders (e.g., Secord et al., 2007). Thus, the data should be interpreted with respect to visual KP rather than KP principles in general. Moreover, it is possible that other types of visual KP (e.g., acoustic visual feedback, electropalatography) might yield different results.

With respect to principles of motor learning (e.g., Maas et al., 2008), the notion that visual KP could enhance acquisition but be detrimental to learning may be time-dependent; that is, visual KP may be particularly beneficial during early stages of practice but if it is continued for too long of a duration, learners may become reliant on the feedback. Therefore, there is likely a point during the learning process in which KP may be most beneficial and a point where it might be most detrimental; indeed, acquisition data from participant 1003 shows the greatest advantage for ultrasound during acquisition in the first few sessions. Further research is needed to identify the optimal sequencing and timing of visual feedback.

5 Conclusion

This study built upon existing research to evaluate the role of US on the acquisition, retention, and generalization of correct production in children with rhotic errors in speech therapy. For some children, acquisition may be facilitated by US visual feedback, and there was no evidence that US visual feedback inhibited retention or generalization. Future research should address the appropriate amount of visual feedback or combination of therapies with and without US, and should continue to explore this approach for different clinical populations (differing age groups and differing profiles of speech disorders). The reasons for individual responses to therapy should be considered, and modifications to the treatment program should be explored to determine how best to facilitate acquisition and learning. At present, US visual feedback continues to be one treatment option for some children with /ɹ/ distortions.

Acknowledgments

This study was completed as part of a masters thesis in fulfillment of a Master of Science degree. Thanks to graduate students Olivia Flaherty, Kerry McNamara, and Monique Uy for assisting with reliability and fidelity. This research was supported by NIH grant R03DC012152 (J. Preston, PI).

References

Adler-Bock M, Bernhardt BM, 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; 16:128–139. [PubMed: 17456891]

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. [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.

Bernhardt B, Bacsfalvi P, Adler-Bock M, Shimizu R, Cheney A, Giesbrecht N, Radanov B. Ultrasound as visual feedback in speech habilitation: Exploring consultative use in rural British Columbia, Canada. Clinical Linguistics & Phonetics. 2008; 22(2):149–162. [PubMed: 18253873]

Bernhardt B, Gick B, Bacsfalvi P, Adler-Bock M. Ultrasound in speech therapy with adolescents and adults. Clinical Linguistics & Phonetics. 2005; 19(6/7):605–617. [PubMed: 16206487]

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. [PubMed: 12858839]

Boersma P, Weenink D. Praat v 5.4.08. 2014 www.praat.org.

Boyce SE. The articulatory phonetics of /r/ for residual speech errors. Seminars in Speech and Language. 2015; 36(4):257–270. [PubMed: 26458201]

Bulte I, Onghena P. An R package for single-case randomization tests. Behavior Research Methods. 2008; 40:467–478. [PubMed: 18522057]

Dollaghan C, Campbell TF. Nonword repetition and child language impairment. Journal of Speech, Language, and Hearing Research. 1998; 41(5):1136–1146.

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

Epstein MA. Ultrasound and the IRB. Clinical Linguistics & Phonetics. 2005; 19(6–7):567–572. [PubMed: 16206484]

Edgington ES. Randomized single-subject experiments and statistical tests. Journal of Counseling Psychology. 1987; 34(4):437.

Gick B, Bacsfalvi P, Bernhardt BM, Oh S, Stolar S, Wilson I. A motor differentiation model for liquid substitutions in children's speech. Proceedings of Meetings on Acoustics. 2007; 1(1):060003.doi: 10.1121/1.2951481

Goldman R, Fristoe M. Goldman Fristoe Test of Articulation. 2. Minneapolis, MN: Pearson; 2000.

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]

Hitchcock ER, McAllister Byun T. Enhancing generalisation in biofeedback intervention using the challenge point framework: A case study. Clinical Linguistics & Phonetics. 2015; 29(1):59–75. [PubMed: 25216375]

Hodges NJ, Franks IM. Learning a coordination skill: Interactive effects of instruction and feedback. Research Quarterly for Exercise and Sport. 2001; 72(2):132–142. [PubMed: 11393876]

Klein HB, McAllister-Byun T, Davidson L, Grigos MI. A multidimensional investigation of children’s /r/ productions: Perceptual, ultrasound, and accoustic measures. American Journal of Speech-Language Pathology. 2013; 22:540–553. [PubMed: 23813195]

Kratochwill TR, Levin JR. Single-Case Intervention Research. Washington, DC: American Psychological Association; 2014.

Lewis BA, Freebairn LA, Taylor HG. Academic outcomes in children with histories of speech sound disorders. Journal of Communication Disorders. 2000; 33(1):11–30. [PubMed: 10665511]

Maas E, Robin DA, Hula SNA, Freedman SE, Wulf G, Ballard KJ, Schmidt RA. Principles of motor learning in treatment of motor speech disorders. American Journal of Speech-Language Pathology. 2008; 17(3):277–298. [PubMed: 18663111]

McAllister-Byun T, Hitchcock ER, Swartz MT. Retroflex versus bunched in treatment for rhotic misarticulation. Evidence from ultrasound biofeedback intervention. Journal of Speech, Language, and Hearing Research. 2014:1–15. [PubMed: 24576834]

Miccio AW. Clinical problem solving assessment of phonological disorders. American Journal of Speech-Language Pathology. 2002; 11(3):221–229.

Newell K, Carlton M, Antoniou A. The interaction of criterion and feedback information in learning a drawing task. Journal of Motor Behavior. 1990; 22(4):536–552. [PubMed: 15117661]

Preston JL, Edwards ML. Phonological processing skills of adolescents with residual speech sound errors. Language, speech, and hearing services in schools. 2007; 38(4):297–308.

Preston JL, Maas E, Whittle J, Leece MC, Mccabe P. Limited acquisition and generalisation of rhotics with ultrasound visual feedback in childhood apraxia. Clinical linguistics & phonetics. 2015:1–17. (ahead-of-print).

Preston JL, McCabe P, Rivera-Campos A, Whittle JL, Landry E, Maas E. Ultrasound visual feedback treatment and practice variability for residual speech sound errors. Journal of Speech, Language, and Hearing Research. 2014

R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2015. URL https://www.R-project.org/

Ruscello D. Visual feedback in treatment of residual phonological disorders. Journal of Communication Disorders. 1995; 28:279–302. [PubMed: 8576411]

Rvachew S. Application of single subject randomization designs to communicative disorders research. Human Communication Canada. 1988; 12(4):7–13.

Rvachew S. Speech perception training can facilitate sound production learning. Journal of Speech Language and Hearing Research. 1994; 37:347–357.

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.

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

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

Shriberg LD. A response evocation program for /ɝ/. Journal of Speech and Hearing Disorders. 1975; 40:92–105. [PubMed: 1123932]

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

Shriberg LD. Childhood speech sound disorders: From postbehaviorism to the postgenomic era. In: Paul R, Flipsen P, editorsSpeech sound disorders in children. San Diego: Plural Publishing; 2009.

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. [PubMed: 20831378]

Shriberg LD, Kwiatkowski J. A follow-up study of children with phonologic disorders of unknown origin. Journal of Speech and Hearing Disorders. 1988; 53(2):144–155. [PubMed: 3361857]

Skelton SL, Hagopian AL. Using tandomized variable practice in the treatment of Childhood Apraxia of Speech. American Journal of Speech-Language Pathology. 2014; 23(4):599–611. [PubMed: 25017177]

Smit AB, Hand L, Freilinger JJ, Bernthal JE, Bird A. The Iowa Articulation norms project and its Nebraska replication. Journal of Speech and Hearing Disorders. 1990; 55(4):779–798. [PubMed: 2232757]

Sjolie G. Masters Thesis. Syracuse, NY: Syracuse University; 2015. Effects of ultrasound as visual feedback of the tongue on generalization, retention, and acquisition in speech therapy for rhotics.

Thoonen G, Maassen B, Gabreels F, Schreuder R. Validity of maximum performance tasks to diagnose motor speech disorders in children. Clinical Linguistics & Phonetics. 1999; 13:1–23.

Thoonen G, Maassen B, Wit J, Gabreels F, Schreuder R. The integrated use of maximum performance tasks in differential diagnostic evaluations among children with motor speech disorders. Clinical Linguistics & Phonetics. 1996; 10(4):311–336.

Volin RA. A relationship between stimulability and the efficacy of visual biofeedback in the training of a respiratory control task. American Journal of Speech-Language Pathology. 1998; 7(1):81–90.

Wagner RK, Torgesen JK, Rashotte CA, Pearson NA. Comprehensive Test of Phonological Processing. 2. Austin, TX: Pro-ed; 2013.

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

Wiig EH, Semel E, Secord WA. Clinical Evaluation of Language Fundamentals. 5. Bloomington, MN: Pearson; 2013.

Wolfe V, Presley C, Mesaris J. The importance of sound identification training in phonological intervention. American Journal of Speech-Language Pathology. 2003; 12:282–288. [PubMed: 12971817]

Appendix A

Sample Probe Lists

Sample US Training Probe	Generalization Probe
/ɑɹ/ arm	/ɚ/ doctor
/ɑɹ/ car	/ɚ/ beautiful flower
/ɑɹ/ far	/ɑɹ/ barn
/ɑɹ/ star	/ɑɹ/ big yard
/ɑɹ/ par	/ɛɹ/ arrow
/ɪɹ/ deer	/ɛɹ/ white hair
/ɪɹ/ near	/ɝ/ birthday
/ɪɹ/ fear	/ɝ/ purple coat
/ɪɹ/ tear	/ɪɹ/ pioneer
/ɪɹ/ gear	/ɪɹ/ steering wheel
	/ɔɹ/ score
Sample NoUS Training Probe	/ɔɹ/ making s'mores
/ɑɹ/ mar	/ɹ/ rice
/ɑɹ/ tar	/ɹ/ red shoes
/ɑɹ/ jar	/gɹ/ sweet grapes
/ɑɹ/ spar	/dɹ/ drinking milk
/ɑɹ/ ark	/fɹ/ frying pan
/ɪɹ/ clear	/pɹ/ pretty sky
/ɪɹ/ cheer	/stɹ/ straw
/ɪɹ/ mere	/spɹ/ yellow sprinkles
/ɪɹ/ hear	/θɹ/ three
/ɪɹ/ peer	/ɹ/ Tina loves rain.
	/ɝ/ I got a fern.
	/ɚ/ Superman is awesome.
	/ɹ/ The room is painted pink.

Appendix B

Sample Training Chains

/ɑɹ/	car	cartoons	old Sunday cartoons	She saw new Monday cartoons.
				Watching funny cartoons.

				Beth hates watching cartoons.
				I laughed at the cartoons.
				The cartoons put us asleep.
				Cartoons begin at six.
/ɪɹ/	gear	headgear	ugly headgear	I do not have headgear.
				His headgear was yellow.
				She got headgear last night.
				Ben has old blue headgear.
				He will not use his headgear.
				He lost his new headgear.
/ɛɹ/	bear	barefoot	walking barefoot	I like being barefoot.
				I always walk barefoot.
				Don’t walk barefoot in the sand.
				Walk barefoot in the house.
				Only kids play barefoot.
				Jake was barefoot at school.
/ɔɹ/	core	encore	standing encore	The encore was too loud.
				They will give one encore.
				Did they play an encore?
				How was the encore tonight?
				John asked for an encore.
				The band gave an encore.

Author Manuscript

Author Manuscript

Author Manuscript

Highlights

- Ultrasound visual feedback during speech therapy may facilitate acquisition of /r/

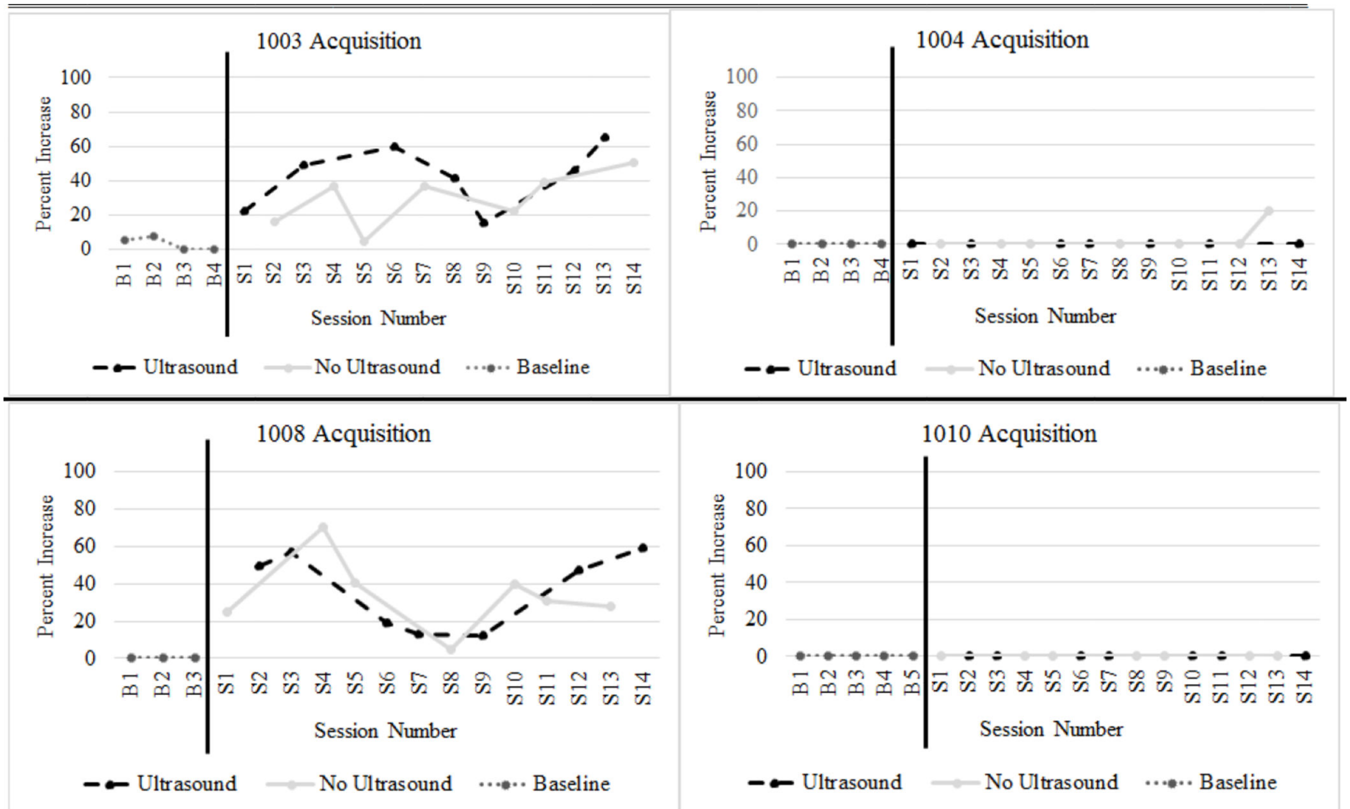

Figure 1. Acquisition scores for participants

*B = Baseline; S = Session

**The baseline for Acquisition used in the above graphs is reported as a percent correct (the average taken from four listeners) of the two rhotic contexts chosen for training. Follow-up data was not able to be measured given the procedures required for measuring acquisition rate.

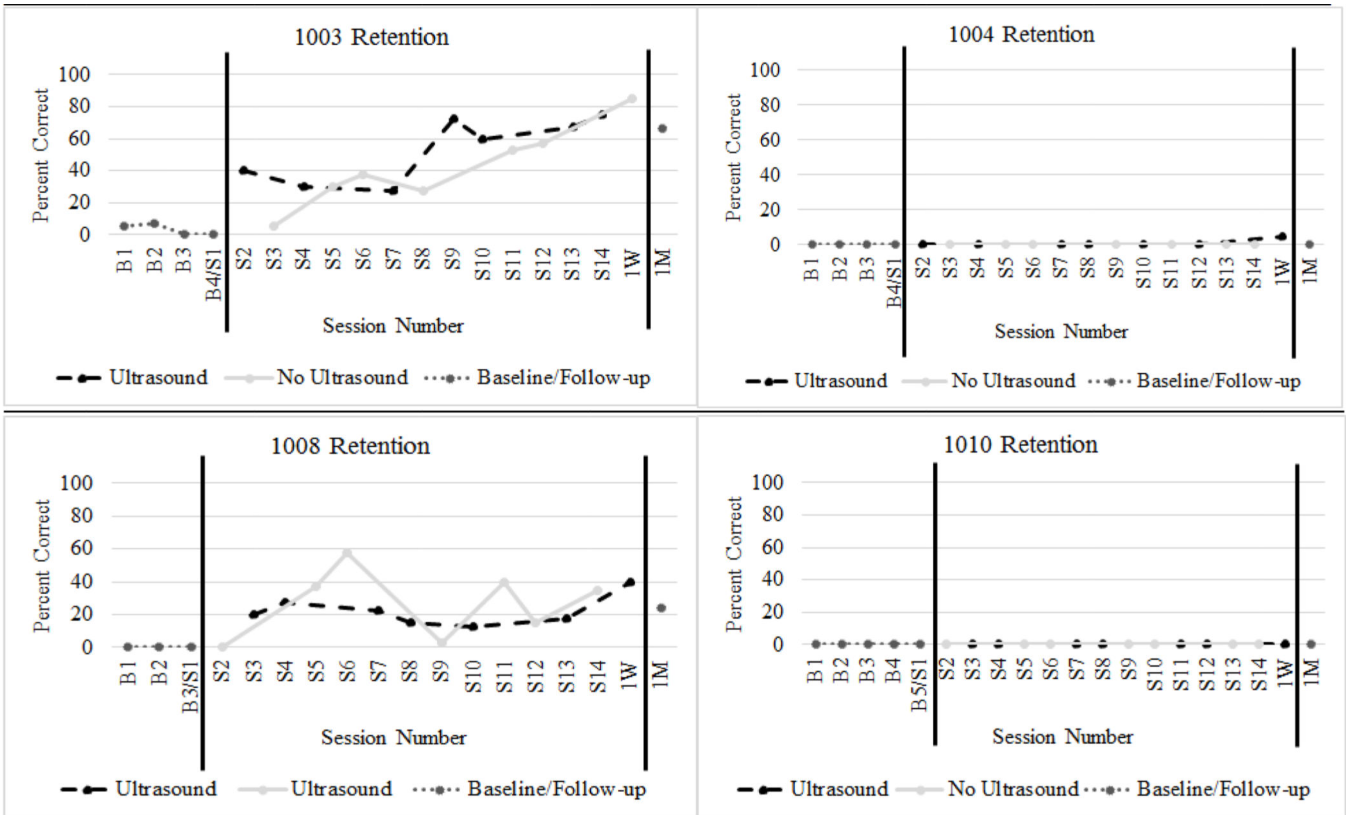

Figure 2. Retention scores for participants

*B = Baseline; S = Session; 1W = 1 week follow-up session; 1M = 1 month follow-up session

**The baseline for retention used in the above graphs is reported as a percent correct (the average taken from four listeners) of the two rhotic contexts chosen for training.

***Follow-up data is reported as a percent correct of all trained items that comprised the last treatment session's Training Probe list (20) that were retained one month after treatment ended.

****The Training Probe list was administered at the beginning of the next consecutive session: thus, session 2 reflects the influence of an ultrasound/no ultrasound session that occurred in session 1.

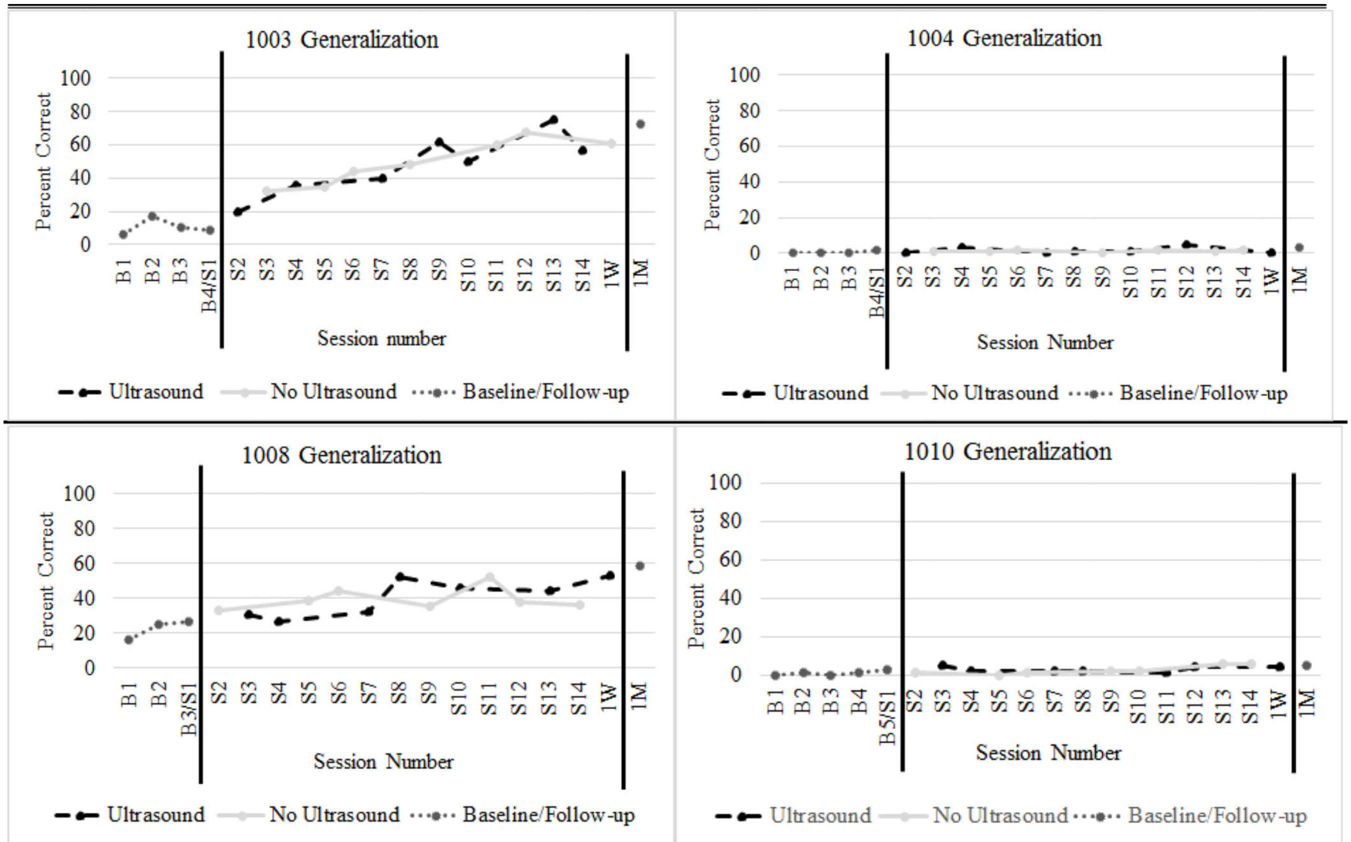

Figure 3. Generalization scores for participants

*B = Baseline; S = Session; 1W = 1 week follow-up session; 1M = 1 month follow-up session

**The baseline and follow-up data for generalization used in the above graphs is reported as a percent correct (the average taken from four listeners) of the Generalization Probe list.

***The Generalization Probe list was administered at the beginning of the next consecutive session: thus, session 2 reflects the influence of an ultrasound/no ultrasound session that occurred in session 1.

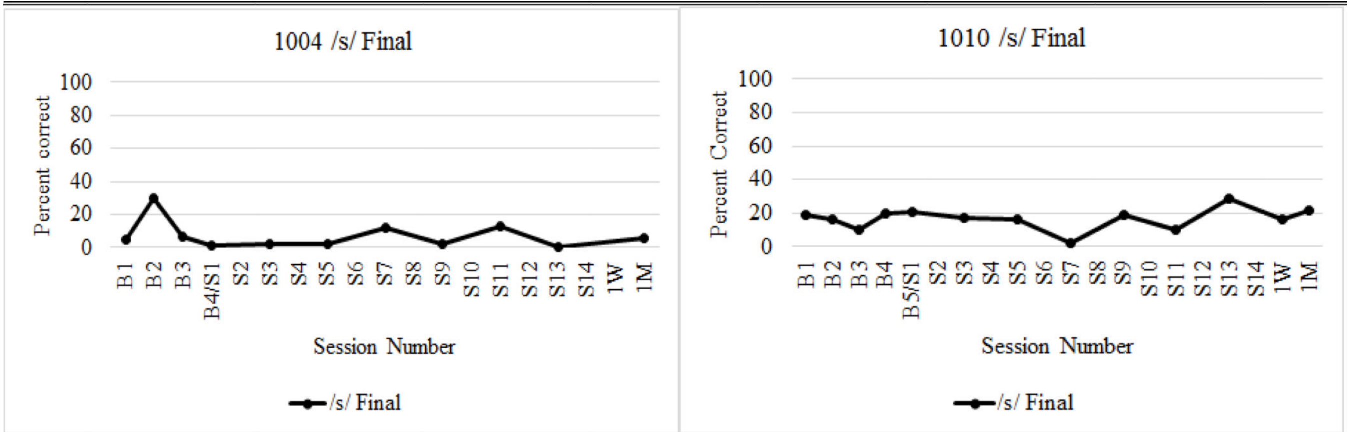

Figure 4. Final /s/ as Untreated Sound Probe

*B = Baseline; S = Session; 1W = 1 week follow-up session; 1M = 1 month follow-up session

**The baseline and follow-up data for the Untreated Sound Probe list used in the above graphs is reported as a percent correct (the average taken from four listeners) of the same probe list administered throughout treatment.

Table 1

Descriptive Data for Participants

	Participants			
	1003	**1004**	**1008**	**1010**
Age (years)	9;0	8;4	9;7	7;0
Intervention History	18 mos-3 yrs; 2nd grade-present	3 yrs-present	7 yrs-present	3 yrs-present
Articulation				
GFTA-2 Standard Score/Percentile	70/3	60/<1	73/1	73/7
Sentence Imitation Task percent /ɹ/ correct	6	0	15	0
Oral Language				
PPVT-4 Standard Score/Percentile	103/58	96/39	111/77	106/66
CELF-4 Formulated Sentences Scaled Score/Percentile	11/63	9/37	7/16	7/16
CELF-4 Recalling Sentences Scaled Score/Percentile	12/75	12/75	8/25	9/37
Phonological Processing				
CTOPP-2 Elision Scaled Score/Percentile	9/37	10/50	5/5	9/37
CTOPP-2 Blending Scaled Score/Percentile	10/50	7/16	2/<1	11/63
Nonword Repetition Percent Phonemes Correct	83	75	82	79
SAILS Number Correct (of 100)	86	84	70	72
SAILS /ɹ/ correct (of 20)	17	19	14	14
Motor Speech				
Multisyllabic Word Repetition PCC	83	75	71	69
Max Performance Task Apraxia Score	0	0	0	2*
Max Performance Task Dysarthria Score	0	0	0	0
Nonverbal				
WASI-2 Matrix Reasoning T-score	55	48	41	47
Stimulability				
Miccio /ɹ/ Stimulability (Max=33)	13	0	18	0

Note: Standard scores are normed with a mean of 100 and standard deviation (SD) of 15. Scaled scores are normed with a mean of 10 and SD of 3. T-scores have a mean of 50 and SD of 10. See Participants section for descriptions of tasks.

*Participant 1010 lengthened the vowel in syllable sequences (e.g., /pʌtəkʌ:: pʌtəkʌ::/), resulting in slow productions and hence an apraxia score of 2.

Table 2

Similarities and differences between US and NoUS Session

Task Order	Duration	US Treatment Session	NoUS Treatment Session
Generalization and Training Probes	4–6 min	Includes US-specific Training words	Includes Non-US-Specific Training words
Untreated Sound Probe	2–3 min (every other session)		
Time Period A	13 min	Therapy with ultrasound	Therapy with no ultrasound
Break	2 min		
Time Period B	13 min	Therapy with no ultrasound	Therapy with no ultrasound
Break	2 min		
Time Period C- I	7 min	Therapy with ultrasound	Therapy with no ultrasound
Time Period C-II	6 min	Random practice with ultrasound	Random practice with no ultrasound
Training Probes	2–3 min		

Author Manuscript

Author Manuscript

Author Manuscript

Table 3

Number of trials (out of six) with verbal feedback in both the US and NoUS Treatment Conditions

Verbal Feedback	Syllable	Mono-syllabic Word	Multi-syllabic Word	Short Phrase Level	Varied Longer Phrase Level	Random Practice
KR verbal feedback	0/6	1/6	2/6	2/6	2/6	2/6
KR and **KP verbal feedback**	5/6	4/6	2/6	1/6	0/6	1/6
Total Trials with Verbal Feedback	5/6	5/6	4/6	3/6	2/6	3/6

Table 4

Acquisition descriptive statistics

	Participants			
Calculation	**1003**	**1004**	**1008**	**1010**
US $\bar{x}$(SD)	42.6 (18.2)	0 (0)	36.8 (21.2)	0 (0)
NoUS $\bar{x}$(SD)	29.4 (15.6)	2.86 (7.56)	34.2 (19.8)	0 (0)
$\bar{x}_{US} - \bar{x}_{NoUS}$	13.2	−2.86	2.67	0
Cohen's d	.783	−.756	.130	N/A
P-value (NoUS–US)	.969	.5	.492	1.0
P-value (US–NoUS)	.039 *	1.0	.516	1.0

$\bar{x}$ = mean percent from four listeners; SD = standard deviation;

* denotes statistical significance

Table 5

Retention descriptive statistics

	Participants			
Calculation	**1003**	**1004**	**1008**	**1010**
US $\bar{x}$(SD)	53.2 (20.3)	0.71 (1.89)	22.1 (9.29)	0 (0)
NoUS $\bar{x}$(SD)	42.1 (25.6)	0 (0)	26.8 (21.4)	0 (0)
$\bar{x}_{US} - \bar{x}_{NoUS}$	11.1	0.71	−4.64	0
Cohen's d	.482	.756	−.303	N/A
P-value (NoUS-US)	.875	1.0	.297	1.0
P-value (US-NoUS)	.172	.5	.711	1.0

$\bar{x}$ = mean percent from four listeners; SD = standard deviation;

*denotes statistical significance

Table 6

Generalization descriptive statistics

	Participants			
Calculation	**1003**	**1004**	**1008**	**1010**
US $\bar{x}$(SD)	48.6 (18.3)	1.43 (1.90)	40.7 (10.6)	2.86 (1.46)
NoUS $\bar{x}$(SD)	49.7 (13.7)	1.29 (0.76)	39.6 (6.46)	2.57 (2.44)
$\bar{x}_{US} - \bar{x}_{NoUS}$	−1.14	0.14	1.08	0.29
Cohen's d	−.071	.107	.127	0.15
P-value (NoUS-US)	.383	.570	.594	.688
P-value(US-NoUS)	.648	.5	.422	.344

$\bar{x}$ = mean percent from four listeners; SD = standard deviation;

*denotes statistical significance

Table 6 note: There was no statistically significant advantage of one treatment over the other in participants' generalization scores.