

## Research Report

# Motor-based treatment with and without ultrasound feedback for residual speech-sound errors

Jonathan L. Preston<sup>†</sup>, Megan C. Leece<sup>‡</sup> and Edwin Maas<sup>§</sup>

<sup>†</sup>Department of Communication Sciences and Disorders, Syracuse University, Syracuse, NY and Haskins Laboratories, New Haven, CT, USA

<sup>‡</sup>Syracuse University, Syracuse, NY, USA

<sup>§</sup>Temple University, Philadelphia, PA, USA

(Received December 2015; accepted March 2016)

### Abstract

*Background:* There is a need to develop effective interventions and to compare the efficacy of different interventions for children with residual speech-sound errors (RSSEs). Rhotics (the r-family of sounds) are frequently in error American English-speaking children with RSSEs and are commonly targeted in treatment. One treatment approach involves the use of ultrasound visual feedback of the tongue.

*Aims:* Although prior studies have shown that children with RSSEs acquire rhotics and generalize to untrained words with ultrasound visual feedback treatment, predictions from schema-based motor learning theory suggest that visual feedback might impede generalization. Therefore, the primary aim was to compare the generalization of rhotics treated with and without ultrasound in children with RSSEs.

*Methods & Procedures:* Twelve children aged 10–16 years with RSSEs affecting rhotics participated in a multiple-baseline single-case design with two treatment phases. For each participant, rhotics in one syllable position were treated for 7 h-long sessions with ultrasound visual feedback and rhotics in a different syllable position were treated without ultrasound in a second treatment phase. The order of treatment conditions was counterbalanced across participants. A treatment framework incorporating the principles of motor learning through chaining procedures was implemented across both treatment phases; thus the primary distinction between conditions was the use of ultrasound visual feedback.

*Outcomes & Results:* On average, both treatments resulted in an approximately 30% increase in accuracy of untreated words in seven sessions. However, variability in response suggested some children showed a preferential response to one treatment over another, some responded well to both interventions, and some responded minimally to both interventions.

*Conclusions & Implications:* Motor-based treatment with and without ultrasound visual feedback of the tongue may aid in speech-sound acquisition for children with RSSEs. Both approaches may be viable options for some children. Future research is necessary to determine which children are the best candidates for interventions with and without ultrasound visual feedback.

*Keywords:* residual speech-sound errors, ultrasound, motor learning, rhotics.

### What this paper adds

Ultrasound visual feedback of the tongue is one option for improving speech-sound accuracy in children with RSSEs. However, no studies have attempted to compare treatment outcomes with and without ultrasound visual feedback. This study addresses this question using a within-participant design. On average, motor-based treatments with and without ultrasound visual feedback resulted in comparable improvements in speech-sound accuracy. However, individual response to treatment suggested that some children benefit more from one form of treatment than from the other.

## Introduction

School-age children with residual speech-sound errors (RSSEs) have phoneme-specific errors that typically affect late-developing sounds. In American English, RSSEs frequently include errors on lingual phonemes such as /ɹ, l, s, z/, although the majority of errors are on /ɹ/ (Shriberg 2009). Although speech-sound disorders are among the most commonly treated communication disorders in children (ASHA 2014), current approaches to treatment are not ubiquitously effective. Historically, the primary approach to treatment has involved verbal instruction in articulator placement for the sound(s) in error, along with practising, feedback and cueing from the speech-language pathologist (SLP). Recent treatment research on this population has explored biofeedback approaches, which involve instrumental applications for real-time visualization of speech. Additionally, applications of schema-based motor learning theory to speech therapy have resulted in the development of treatment procedures that are structured to incorporate principles of motor learning (PML) (Hitchcock and McAllister Byun 2015, Preston *et al.* 2014). To date, no studies have compared outcomes of PML-based interventions with and without visual feedback for children with RSSEs. The present investigation sought to compare PML-based interventions that do and do not include visual feedback of the tongue for /ɹ/ distortions in children with RSSEs.

### *Articulatory requirements of /ɹ/*

Two primary constrictions are formed with the tongue during a typical production of /ɹ/: an oral constriction formed by elevation of the front of the tongue (tip, blade or anterior dorsum) and a pharyngeal constriction formed by posterior movement of the tongue root. Moreover, the lateral margins of the tongue are typically elevated toward the molars, creating a groove in the midline of the tongue. The fact that multiple lingual constrictions are required may be one reason why /ɹ/ can be difficult for some children to learn (Boyce 2015, Magloughlin 2016). In addition, a variety of different tongue shapes can be used to achieve these oral and pharyngeal constrictions (including tongue shapes that are classically described as ‘bunched’ and ‘retroflex’) and there may be variability in tongue shapes from person to person and from context to context (Delattre and Freeman 1968). The complexity and the variability in tongue shape may contribute to the difficulty in teaching /ɹ/ in speech therapy (Boyce 2015).

### *Motor-based approaches for /ɹ/ remediation*

Traditional approaches for remediating sound errors associated with RSSEs include individual or small group

instruction. The focus is typically on eliciting an acceptable version of the phoneme in isolation or in a particular phonetic context, followed by systematic practice to stabilize productions. A hierarchy of linguistic complexity is typically followed that advances to syllables, words, phrases, sentences and structured conversation once the child has reached pre-established criterion (e.g., 80% or 90% accuracy) in a session. Verbal feedback on articulatory position is typically provided by the SLP following an error and positive reinforcement is typically provided following correct productions.

A number of PML, derived from the non-speech motor learning literature, can be applied to speech-sound intervention. These principles dictate elements of treatment that can impact acquisition and learning (Maas *et al.* 2008). *Acquisition* refers to the learner’s performance during practice, whereas *motor learning* is captured by changes in performance on retention and generalization tasks (maintenance of improvement over time and changes to untrained tasks or words, respectively). The important principles that influence acquisition and learning include aspects of how speech is practised and how feedback is provided. For example, practising simple targets (e.g., ‘ray’) may yield high performance during acquisition but might result in less learning than practising on more complex targets (e.g., ‘two straight lines’); practising on complex targets would be more likely to lead to retention and generalization. The amount of practice (i.e., dose) can also contribute to acquisition and learning (e.g., Williams 2012). Given the importance of lasting and widespread improvements, our focus is on measures of learning, i.e., retention and generalization performance (Maas *et al.* 2008).

Critical to the present investigation, feedback is believed to play an important role in motor learning (e.g., Preston *et al.* 2014). In speech therapy, feedback can take multiple forms: (1) no feedback, (2) knowledge of results (KR) feedback, which is feedback on accuracy of a production as correct or incorrect, and (3) knowledge of performance (KP) feedback, which is feedback on the nature of the movement (Maas *et al.* 2008, Preston *et al.* 2014). In traditional speech therapy, KP feedback is provided verbally (e.g., ‘You didn’t raise the front of your tongue when you said /ɹ/’). However, KP can also be provided visually with instrumental methods such as ultrasound that display essential elements of the target speech movement. For children who are having difficulty acquiring /ɹ/, KP in the form of ultrasound might aid acquisition of movements for which they do not have an accurate internal model or representation. This first hypothesis (KP is beneficial for learning complex movements) is consistent with some evidence from the motor learning literature which suggests that KP feedback compared with KR feedback may facilitate the learning of complex movement patterns for which the

task goal may not be immediately clear to the learner (Newell *et al.* 1990). Given that /ɪ/ is an articulatorily complex sound for which children may not have a clear idea of the movement goals, visual feedback about critical movement patterns is expected to facilitate retention and generalization compared with treatment without visual feedback. However, there are also suggestions in the motor learning literature that too much KP feedback might hinder learning (Hodges and Franks 2001), even though such feedback may facilitate acquisition. Thus, according to this alternative hypothesis (KP impedes learning), visual KP feedback may facilitate acquisition but hinder retention and generalization relative to treatment without visual feedback. Possible explanations include the notion that learners become overly dependent on external feedback, and the notion that the addition of visual feedback increases cognitive processing demands related to integrating feedback with self-generated appraisal of the quality and success of the movement (e.g., Hodges and Franks 2001). In summary, two different hypotheses derived from the motor-learning literature make opposing predictions regarding the effects of visual KP on learning. The present study was designed to test these predictions with children with RSSEs using ultrasound biofeedback as the means for providing visual KP.

#### *Visual feedback in speech-sound intervention*

Visual feedback approaches to therapy involve the use of instrumentation to display crucial aspects of speech. For example, electropalatography and spectral visual feedback are approaches that incorporate visual feedback to teach productions of target speech sounds. Ultrasound has recently been reported as another option for visual feedback, as it provides information about tongue shape; knowing about intended and executed tongue movements may help children with RSSEs to achieve correct articulation of /ɪ/. Therefore the present investigation utilizes ultrasound, which shows the shape of the tongue and allows visualization from tip to root (in sagittal view) and from side to side (in coronal view).

Ultrasound uses reflected high-frequency sound waves to generate images in real time. By holding the ultrasound transducer beneath the chin, the upper contour of the tongue can be visualized creating an opportunity to cue desired articulatory movements and to compare intended movements with actual movements. For /ɪ/, sagittal views with ultrasound can show movements of the tongue that are responsible for both the oral and pharyngeal constrictions. This is particularly advantageous as it allows for visualization of essential movements for /ɪ/ that are not at all visible during traditional treatment. Additionally, coronal views can be used to visualize the left-to-right dimension which can reveal elevation of the

lateral margins of the tongue that may be associated with an acceptable /ɪ/ production.

Common articulatory correlates of distorted /ɪ/ that may be visible with sagittal ultrasound views include a constriction with the tongue dorsum in the velar region, a low tongue tip/blade and a lack of posterior movement of the tongue root (e.g., Adler-Bock *et al.* 2007, Preston *et al.* 2016); in coronal views, distorted productions of /ɪ/ may lack elevation of the sides of the tongue. Therefore, cueing with an ultrasound may include instructions to elevate the anterior tongue, lower the dorsum, retract the root and raise the sides of the tongue. However, the primary goal is to find a tongue shape that results in an acoustically acceptable production, not to achieve a predetermined tongue shape (McAllister Byun *et al.* 2014).

To date, a number of studies have explored the use of ultrasound visual feedback of the tongue to improve production of a number of lingual phonemes in individuals with RSSEs and other speech disorders. With respect to /ɪ/, case studies and single-subject experimental designs have demonstrated that some (though not necessarily all) individuals with RSSEs who undergo speech-sound training with ultrasound visual feedback improve their productions of trained and untrained words. For example, Adler-Bock *et al.* (2007) reported on two participants, aged 12 and 14 years, who showed improvement in /ɪ/ at the phrase level following 14 treatment sessions with ultrasound visual feedback (improvements of 48% and 62% respectively). McAllister Byun *et al.* (2014) reported two single-case design studies; treatment response was minimal in the first study when participants were required to produce a 'bunched' /ɪ/ tongue shape; in the second study when participants were encouraged to explore different tongue shapes for /ɪ/ the mean improvement of consonantal and vocalic /ɪ/ was 54% in 17 sessions. Preston *et al.* (2014) reported outcomes from eight participants aged 10–20 years who were treated on /ɪ/ (and other phonemes) with ultrasound visual feedback, with mean improvement on untrained words of approximately 35% in seven sessions. Other studies have also shown improvements on /ɪ/ productions in children with childhood apraxia of speech (Preston *et al.* 2013), although some may fail to acquire /ɪ/ during treatment and/or to generalize (Preston *et al.* 2016). Research has also demonstrated improvements on /ɪ/ and for adolescents with hearing impairment (e.g., Bacsfalvi 2010). Although treatment response varies across individuals and across studies, there is evidence that some individuals can improve /ɪ/ accuracy with treatments that include ultrasound visual KP feedback; this approach may therefore serve as an alternative or a supplement to traditional therapy for individuals whose errors have not responded to standard interventions. This would be consistent with the KP-is-beneficial

hypothesis. However, as described above, according to the KP-impedes-learning hypothesis, also derived from the motor learning literature, generalization may be hindered by overreliance on KP feedback as may be provided with ultrasound (Maas *et al.* 2008).

### *Purpose and hypotheses*

Treatments for RSSEs that include visual feedback of the tongue have been shown to improve speech-sound accuracy for /ɪ/ distortions. However, because some children may improve /ɪ/ accuracy with structured motor-based treatment that does not include visual feedback, it is important to examine if treatment response is better with ultrasound feedback than without. No studies to date have directly compared the effects of treatment with versus without ultrasound while keeping other aspects of treatment the same. The aim of the study was to provide a preliminary investigation of response to motor-based treatment with and without ultrasound visual feedback. A within-participant design replicated across participants was conducted to explore whether ultrasound KP feedback might facilitate or inhibit learning (retention and generalization of improvements on production of rhotics). According to hypothesis 1 (KP is beneficial for learning), a PML-based intervention with ultrasound should result in greater learning than a PML-based intervention without ultrasound, possibly by providing learners with explicit information about how to solve a complex pattern of speech movements. By contrast, according to hypothesis 2 (KP impedes learning), visual feedback is predicted to impede learning, possibly due to the cognitive processing demands associated with integrating such feedback and/or due to overreliance on such feedback. Therefore, we sought to explore these competing hypotheses by evaluating whether PML-based intervention with or without ultrasound will lead to better retention and generalization.

## **Methods**

### *Participants*

Twelve children aged 10–16 years (eight male, four female) with RSSEs affecting /ɪ/ were recruited by referrals from local SLPs. Six children were recruited in southern Connecticut and an additional six children were recruited in the greater Syracuse, New York area (both are regions with a rhotic American English dialect). None had diagnosed developmental disabilities or cognitive delays that might impact speech (e.g., autism, Down syndrome). Intervention histories varied (table 1), but all participants were diagnosed with a speech-sound disorder and continued to have distorted /ɪ/ productions despite previous treatment. At the time of the study, none was working on /ɪ/ with another SLP because

they were (1) no longer eligible for speech–language services, (2) enrolled in the study during summer and were not receiving school-based therapy, (3) working with an SLP on a speech or language goal other than /ɪ/, or (4) on a waiting list to begin receiving private therapy services.

Speech-sound accuracy was assessed with a variety of tasks. To confirm the presence of a RSSE, children had to score at or below the fifth percentile on the Goldman–Fristoe Test of Articulation—2 (Goldman and Fristoe 2000). To confirm the participants had errors on rhotics, they were required to score below 25% accuracy on word probes evaluating /ɪ/ (see generalization probes below). These probes were used to determine eligibility for the study and also to track progress over the course of therapy. Additionally, a sentence imitation task sampled rhotic accuracy in 53 words in 15 sentences (e.g., ‘The cops arrest a robber at the drug store’). Data from pre-treatment assessments are presented in table 1.

All participants passed a pure tone hearing screening bilaterally at 25 dB at 500, 1000, 2000 and 4000 Hz. Participants were required to have receptive language skills broadly within the average range as defined by a standard score above 80 on the Peabody Picture Vocabulary Test—4 (Dunn and Dunn 2007).

### *Descriptive measures*

In addition to the above inclusionary criteria, further descriptive data were collected, as such information may provide insight into characteristics of responders or non-responders to treatment. Oral language measures included the Recalling Sentences and Formulated Sentences subtests of the Clinical Evaluation of Language Fundamentals—4 (Semel *et al.* 2003) and the Expressive Vocabulary Test—2 (Williams 2007). Descriptive data on nonverbal IQ were collected using the Wechsler Abbreviated Scales of Intelligence—2 (Wechsler 2011) Matrix Reasoning subtest.

Several measures were used to provide descriptive information on phonological processing, as this domain may be related to RSSEs in some children (e.g., Preston *et al.* 2015). Phonological awareness was evaluated using the Elision, Phoneme Blending and Phoneme Isolation subtests of the Comprehensive Test of Phonological Processing—2, which can be combined to derive a Phonological Awareness Composite score (Wagner *et al.* 2013). Phonological working memory was evaluated by per cent phonemes correct a nonword repetition task that did not include rhotic phonemes (Dollaghan and Campbell 1998). Finally, children’s recognition of speech-sound errors was evaluated using Speech Assessment and Interactive Learning (SAILS) (Rvachew 1994), which included 20 trials each of the phonemes /ʃ, s, f, θ, ɪ/.

**Table 1. Participant characteristics and performance on standardized and non-standardized assessments prior to therapy**

|   | A       | B            | C    | D    | E    | F     | G    | H    | I    | J    | K    | L       |
|---|---------|--------------|------|------|------|-------|------|------|------|------|------|---------|
| Age (years;months)                        | 11;9    | 11;0         | 16;7 | 12;4 | 13;3 | 15;9  | 12;5 | 10;3 | 10;3 | 10;1 | 10;1 | 10;7    |
| Gender                                    | F       | M            | M    | M    | M    | F     | F    | M    | M    | M    | M    | F       |
| History of previous therapy (ages)        | 7, 9–11 | Not reported | 3–13 | 9–12 | None | 13–14 | 5–12 | None | 3–9  | None | 2–9  | 3–6, 10 |
| GFTA-2 standard score                     | 75      | 72           | 41   | 78   | 88   | 82    | 61   | 60   | 78   | 81   | 57   | 77      |
| GFTA-2 percentile rank                    | < 1     | < 1          | < 1  | 3    | 4    | 1     | < 1  | 1    | 4    | 4    | 1    | 1       |
| Percent /ɪ/ correct 15-sentences imitated | 9       | 42           | 13   | 28   | 6    | 38    | 6    | 6    | 13   | 17   | 13   | 0       |
| PPVT-4 standard score                     | 125     | 82           | 99   | 121  | 125  | 109   | 92   | 110  | 108  | 109  | 129  | 100     |
| CELF-4 FS scaled score                    | 9       | 9            | 13   | 15   | 12   | 13    | 1    | 13   | 7    | 13   | 15   | 11      |
| CELF-4 RS scaled score                    | 7       | 7            | 7    | 10   | 14   | 11    | 2    | 10   | 10   | 11   | 13   | 10      |
| EVT-2 standard score                      | 117     | 97           | 92   | 113  | 113  | 110   | 80   | 112  | 90   | 112  | 115  | 100     |
| CTOPP-2 PA composite                      | 100     | 98           | 84   | 92   | 107  | 112   | 84   | 82   | 84   | 100  | 88   | 84      |
| Nonword repetition, % phonemes correct    | 89      | 89           | 84   | 94   | 95   | 92    | 89   | 93   | 88   | 94   | 92   | 69      |
| SAILS % correct (of 100)                  | 90      | 80           | 84   | 87   | 86   | 90    | 86   | 89   | 87   | 79   | 81   | 86      |
| SAILS /ɪ/ correct (of 20)                 | 19      | 14           | 17   | 18   | 16   | 17    | 18   | 16   | 19   | 19   | 14   | 16      |
| MaxPT apraxia score                       | 0       | 0            | 0    | 0    | 0    | 0     | 0    | 0    | 0    | 0    | 0    | 0       |
| MaxPT dysarthria score                    | 0       | 0            | 0    | 0    | 0    | 0     | 0    | 0    | 0    | 0    | 0    | 0       |

Note: GFTA-2 = Goldman-Fristoe Test of Articulation—2; PPVT-4 = Peabody Picture Vocabulary Test—4; EVT—2 = Expressive Vocabulary Test—2; CELF-4 = Clinical Evaluation of Language Fundamentals—4; FS = formulated sentences; RS = recalling sentences; CTOPP = Comprehensive Test of Phonological Processing. Standard scores have a mean of 100 and SD of 15. Scaled scores have a mean of 10 and SD of 3. MaxPT = maximum performance tasks.

To evaluate speech motor function, a maximum performance task was administered (Thoonen *et al.* 1999). This task involved sustained productions of /a/, /f/, /s/, /z/ as well as rapid and repeated productions of /pɑ/, /tɑ/, /ka/, /ma/ and /patakɑ/. Based on the duration of sustained productions and rate and accuracy of repeated productions, all the participants were classified as ‘not dysarthric’ and ‘not apraxic’ on both the dysarthria and apraxia scores (cf. Thoonen *et al.* 1999). Additionally, none of the participants had a pre-existing diagnosis of apraxia of speech or dysarthria.

#### *Intervention design*

An ABACA/ACABA single-case design with multiple baseline across behaviours and participants was used to compare response to treatment in two different treat-

ment phases: PML with ultrasound (PML + US) and without ultrasound (PML + NoUS). As participants entered the study, they were paired and one member of the pair was randomly assigned to begin the first phase of treatment in the PML + US condition while the other member of the pair was assigned to begin the first phase of treatment in the PML + NoUS condition. Each participant then was exposed to the other treatment in a second phase. This ensured an equal representation of participants who began in each condition. For each participant, two different syllable positions were treated—one in each treatment phase (these were selected among onset singleton, onset cluster, nucleus, coda). Therefore, the two syllable positions that were the lowest accuracy for a participant were treated, and these two syllable positions were assigned to the two different conditions.

Baseline generalization probes were collected during assessment sessions before treatment began for three to five data points at a rate of approximately twice per week. Additionally, generalization probes were collected during every other treatment session. Three to five generalization probes were collected between phases I and II, and again after phase II. Each phase of treatment was seven sessions. Finally, two months after the final post-treatment generalization probe, participants returned to determine whether their speech accuracy continued to improve or regressed.

### Generalization probes

Four generalization probe lists were developed that included words with rhotics in various syllable positions: onset singleton /ɪV/ (25 words), onset cluster /Cɪ/ (50 words), nucleus /ɜ:/ (25 words) and coda singleton /Vɪ/ (25 words). Probes were administered by having the participant read the word list; responses were audio recorded for scoring by multiple listeners blind to treatment status. Each word was scored as correct or incorrect based on the quality of the rhotic, and a per cent accuracy was calculated. Items on these lists were not treated in therapy. Generalization probes were administered at the start of sessions and therefore reflect retention of skills following the previous session.

For each child, pre-treatment generalization probe scores were used to identify the two rhotic syllable positions with the lowest per cent accuracy. One syllable position was assigned randomly to each treatment condition; for 10 children targets included one onset (singleton or cluster) and one rime (nucleus or coda). These were selected because onsets and rimes have been shown to emerge relatively independently in children's speech-sound development (McGowan *et al.* 2004) and because prior intervention studies have shown generalization within onset positions, but not between onset and rimes (McAllister Byun and Hitchcock 2012, Preston *et al.* 2013, 2014). For two children for whom onset singletons /ɪV/ and clusters /Cɪ/ were above 30% accurate, the targets selected for treatment included a rhotic nucleus /ɜ:C/ and a rhotic coda (/Vɪ/).

### Intervention procedures

Treatment was scheduled twice per week for 60-min sessions. Treatment was delivered by an ASHA-certified SLP. For each participant, the two treatment targets were randomly assigned to a treatment phase: one to the PML + US phase and the other to the PML + NoUS phase. The two treatment conditions were designed to be identical with the exception of the use of ultrasound visual feedback. For six of the participants, the first phase of treatment was PML + US and the second phase was

PML + NoUS; for the remaining six participants, the order was reversed.

Target words in a session were chosen for a given syllable position by selecting words that contained one of two adjacent phonemes. For example, onset singleton /ɪV/ was trained with two following vowels (e.g., /iɪ, ɔɔ/); onset clusters /Cɪ/ were trained with two different consonants (e.g., /tɪ, bɪ/); coda /Vɪ/ was trained in two vowel contexts (e.g., /ɑɪ, ɔɪ/), and nucleus /ɜ:/ was trained with two consonant codas (e.g., /ɜ:l/, /ɜ:n/). This was done to control for variability of target word selection.

### Session structure

Each 1 h-long session began with elicitation (pre-practice) which focused on achieving six correct renditions of each of the two syllables for the session (e.g., six each of /iɪ, ɔɔ/ for onset /ɪV/, or 12 total). Elicitation was relatively unstructured and included a variety of strategies to facilitate correct production such as imitation, shaping (e.g., shaping /ɪ/ from /ɑ/, /l/, /i/, /f/) and phonetic placement cues presented via pictures and verbal instruction (e.g., pull the back of your tongue back into your throat; lift the front of the tongue up; feel the sides of the tongue touch the insides of the back teeth). If the criterion of 12 correct attempts was not achieved, the entire session was spent in elicitation. If the child was readily stimulative, the elicitation stage could be completed in approximately 1–2 min.

Once the participant achieved 12 correct attempts in the session, treatment progressed from the elicitation stage to structured practice. Structured practice included chaining procedures to build increasingly complex productions from syllables to monosyllabic words, multisyllabic words, phrases, and self-generated sentences based on performance. Within each of these levels, participants practised six attempts at the target item (a syllable, word, phrase or self-generated sentence). If they achieved five of six correct attempts, the next practice item was at the next highest level in the hierarchy. If they achieved fewer than five of six attempts, the next practice item returned to the syllable level. Each series of levels was developed as a chain, with subsequent levels building on the core syllable or word from the previous levels. For example, a participant might practise /iɪ/ and, based on performance, progress to 'read', 'reading', 'reading some books' and a self-generated sentence with the word 'reading'. A chaining example for /ɑɪ/ included /ɑɪ/, 'car', 'cartoon', 'funny cartoon' and self-generated sentences with 'cartoon'.

In structured practice, each level of complexity also had a predetermined proportion of trials with no verbal feedback, knowledge of results (KR) feedback and verbal KP feedback. Verbal feedback was provided on five of

six trials at the syllable level (five with both KP + KR), five of six trials at the monosyllabic word level (two KR, three KP + KR), four of six trials at the multisyllabic word level (two KR, two KP + KR), three of six trials at the phrase level (one KR, two KP + KR), and three of six trials at the self-generated sentence level (one KR, two KP + KR).

### *Condition differences*

Sessions were divided into four 13-min time periods (A–D). In the PML + US condition ultrasound feedback was provided during periods A and C, and periods B and D were conducted without ultrasound. In the PML + NoUS condition, each of the 13-min time periods involved no ultrasound. Thus, the PML + US condition involved 50% ultrasound visual feedback and the PML + NoUS involved 0% ultrasound. The 13-min periods were controlled by use of a timer. Intervention procedures for elicitation and structured practice that are described above occurred during these time periods.

When ultrasound was used in the PML + US condition, visual feedback of the tongue was provided with either a Seemore PI 7.5 MHz ultrasound transducer connected to a Dell personal computer, or an Echo Blaster 128 ultrasound with a PV 6.5 transducer connected to a Dell personal computer. It was generally the case that during the first session the clinician held the transducer beneath the participant's chin to obtain clear images; however, during subsequent sessions the child held the transducer and instructions were provided on positioning of the transducer if necessary. The initial session with the ultrasound involved orienting the participant to the images and children were required to point to the tongue tip, body and root of their tongue on the image. Sagittal views of the tongue were used initially to focus on achieving oral and pharyngeal constrictions, although coronal views were used at the clinician's discretion to address elevation of the lateral margins of the tongue. Cues and feedback in the sagittal view typically focused on encouraging elevating the anterior tongue, lowering the tongue dorsum and retracting the tongue root into the pharynx during /ɪ/. In coronal view, cues addressed raising the lateral margins of the tongue while lowering the midline of the tongue to form a groove.

Cues for /ɪ/ were dependent upon the nature of the error observed and the nature of the target tongue shape that the clinician was attempting to facilitate. The typical pattern of error for /ɪ/ was a high back tongue position of the tongue dorsum with the tongue tip/blade low; additionally, errors typically involved a lack of tongue root retraction. Thus, clinicians were free to cue whatever elements they believed would be helpful, such as raising the front of the tongue, lowering the dorsum or retracting the tongue root into the pharynx. The rela-

tive amount of cueing of any particular movement was dependent upon the child's error and the clinician's judgment of success. Examples of sagittal views of correct and distorted /ɪ/ shapes for three participants are shown in figure 1. In general, the common error for distorted /ɪ/ productions included a high back tongue dorsum with the anterior tongue very low (left column). Correct productions were generally associated with elevation of the anterior tongue and depression of the tongue dorsum indicative of tongue root retraction (right column).

When in the PML + NoUS condition, or in periods B and D of the PML + US condition, verbal cues often addressed similar movement targets but without any visual display. The SLP used similar shaping techniques and phonetic placement cues during elicitation, and provided feedback during structured practice based on an assumption that the child needed to achieve pharyngeal constriction, a low tongue body, an elevation of the anterior or middle tongue, and/or elevation of the lateral margins of the tongue. Thus, as in traditional speech therapy, the SLP essentially 'guessed' based on listening or watching the mouth to determine aspects of /ɪ/ production that might be cued.

### *Session example*

An example of the progression of a session is as follows (cf. Preston *et al.* 2014). Each session began with elicitation to facilitate correct productions of the target syllables. For some children, substantial cueing was required to facilitate correct productions. If a child in the PML + US condition took 30 min to meet the pre-established criterion of 12 correct productions (e.g., six /iɪ/ and six /ɪo/), this would have encompassed 13 min of elicitation with the ultrasound (time period A), 13 min of elicitation without the ultrasound (time period B), and 4 min of elicitation again with the ultrasound (time period C). At this point, the performance criteria to complete elicitation and transition to structured practice would be met, and the participant would begin structured practice at the syllable level. In this example, there would be 9 min remaining in time period C for structured practice with the ultrasound, and the final 13 min of structured practice would occur without the ultrasound (time period D).

During the 9 min of structured practice in time period C with the ultrasound and the 13 min of structured practice in time period D without the ultrasound, the chaining procedures described above would be implemented. Structured practice would begin with six attempts at the first target syllable (e.g., /iɪ/), at which point the clinician made the decision of whether to advance to the word level (if at least five of six were correct, moving to a monosyllabic word such as 'read'), or stay at the syllable level, moving to six attempts of

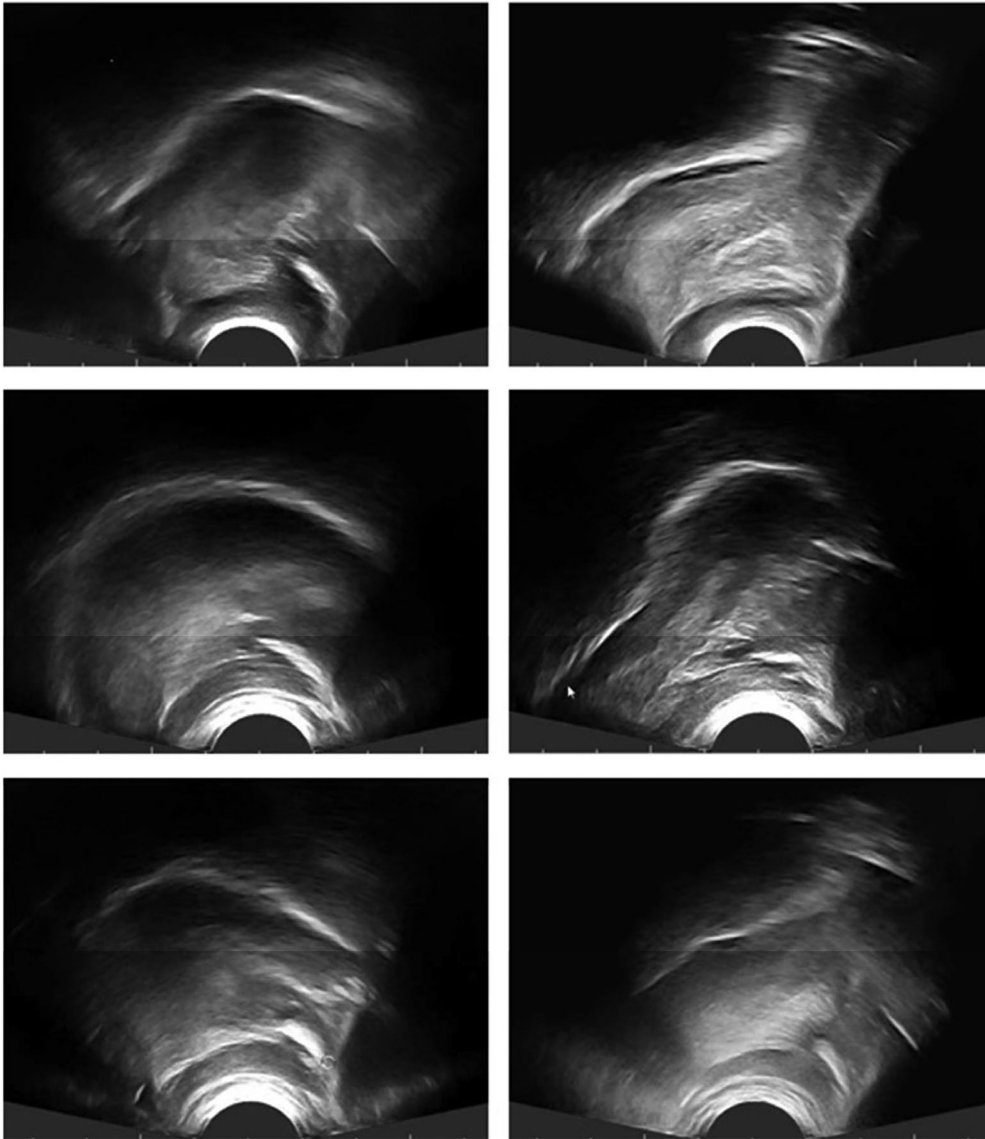


Figure 1. Sample sagittal views of tongue shapes for /ɪ/ before treatment (left column) and after treatment (right column) for three participants. Right is anterior and left is posterior.

/ɪo/. Again, a decision point would be reached after six attempts to determine if the next target would be a higher level of complexity or would return to the syllable level.

#### *Intervention fidelity*

Sessions were audio and video recorded. Screen-capture software was used to synchronize the participant's audio along with video images of the ultrasound, and a back-up audio recording was collected as well. To verify that the specified type and frequency of feedback was delivered, a research assistant reviewed recordings of two randomly selected sessions per participant (24 total sessions). The specified verbal feedback (KP + KR, KR only or no

feedback) was provided, on average, 98% of the time in the sessions that were reviewed (SD = 1.9%, range = 90.6–100%).

As treatment is predicated on the treating clinician's immediate decisions about acoustic acceptability of the participant's productions, the research assistant calculated agreement with the treating clinician's judgment of accuracy (verbal KR). This inter-rater reliability of KR judgment was computed for two sessions per participant. Trials in structured practice were judged for reliability only when KR or KR + KP feedback was given (trials without feedback were not calculated in the overall per cent 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, 94% of the time in the sessions that were reviewed (SD = 4.3%, range = 86.8–100%).

#### *Reliability of generalization probe scores*

All generalization probes on which effect sizes were calculated (pre-treatment baseline, between phases, and post-treatment) were independently scored by a minimum of three listeners who were blind to treatment status; probes administered during treatment phases were used only for graphical display of the data and were scored by a minimum of two listeners, with a third listener required if the first two disagreed by 20% or more (5% of probes). The average absolute difference between each pair of listeners across all probes was 13.2%. Data presented are based on the average score among all listeners for a given probe.

#### *Reliability of the sentence imitation task*

The 15-sentence imitation task was also used to track change from pre- to post-treatment. Two listeners independently scored the sentence imitation task. The mean difference between the two listeners' scores was 8.1% (SD = 6.6%).

#### *Data analysis*

The primary outcome variable was per cent accuracy on the generalization probes. Calculation of per cent increase on probe scores was derived from the probes administered prior to and after the treatment phase. Additionally,  $d_2$ , which represents the per cent increase divided by the pooled pre- and post-treatment standard deviations, was used to compare conditions. Graphical displays of the data were used to evaluate overall trends, as well as to explore continued growth at the 2-month follow-up. We also explored further evidence of generalization and retention through comparison of per cent rhotics correct on the 15-sentence imitation task.

### **Results**

Graphical displays of per cent rhotics correct on generalization probes are shown in figure 2 for participants who were treated with PML + US followed by PML + NoUS, and in figure 3 for participants who were treated with PML + NoUS followed by PML + US. Baseline data were stable for the first sound target that was treated for each of the participants except B. However, for two participants (i.e., A and B) there was some evidence of generalization from one syllable position to another; data from the second phase of treatment for these two participants should be interpreted with caution.

The primary comparison of interest was the difference in improvement on retention/generalization probes for PML + US versus PML + NoUS conditions. Figures 2 and 3 show that most participants improved their accuracy of the treated rhotic in both treatment conditions, although the magnitude of response varied. Table 2 lists the treated rhotic targets in each condition for each participant. Results were explored in terms of raw per cent increase in rhotic accuracy and standardized effect sizes ( $d_2$ ) derived from the four to five generalization probes administered immediately before and immediately following the respective treatment phase. It should be noted that some participants showed continued improvement (i.e., phase I targets continued to improve during phase II); thus, the data in table 2 may underestimate the long-term impact of the interventions. At the group level, a direct comparison could be made between the first phases of treatment. The six participants who were assigned to PML + US in phase I improved their productions by 6.7–74.9% (mean = 31.1%) over the first seven sessions. Participants who were assigned PML + NoUS in phase I improved their productions by 0–72.4% (mean = 30.2%) in phase I. The average increase from before to after each treatment phase was between 28% and 34% for both phase I and phase II, as well as for both PML + US and PML + NoUS. This suggests relatively similar progress across the two treatment conditions.

As an additional way to explore the data, we used a benchmark of 15% improvement to qualify treatment response. This benchmark was selected because it was outside the realm of measurement error and therefore reflected reliable change that was at least half of the mean treatment response (i.e., 30% as indicated above). Seven of 12 participants increased their accuracy on the target rhotic by  $\geq 15\%$  in both conditions (i.e., participants A, B, D, E and G–I). Only two of the 12 participants showed an increase of  $\leq 15\%$  in both of the treatment conditions (i.e., participants K and L), suggesting a general failure to respond to PML-based treatment. Three participants showed an increase above baseline of  $\geq 15\%$  in one condition but not in both (i.e., participants C, F and J).

Comparing within-participant response to the two treatment conditions, we also explored whether any children showed an advantage of  $> 15\%$  for one treatment condition over the other. Four participants showed an advantage ( $> 15\%$  difference) for PML + US over PML + NoUS (i.e., participants A–C and J); three participants showed an advantage for PML + NoUS over PML + US (i.e., participants E–G). The remaining five participants showed gains that were relatively similar between the two treatment conditions (i.e.,  $< 15\%$  difference; participants D, H, I, K and L). Despite individual differences in treatment response, on average there

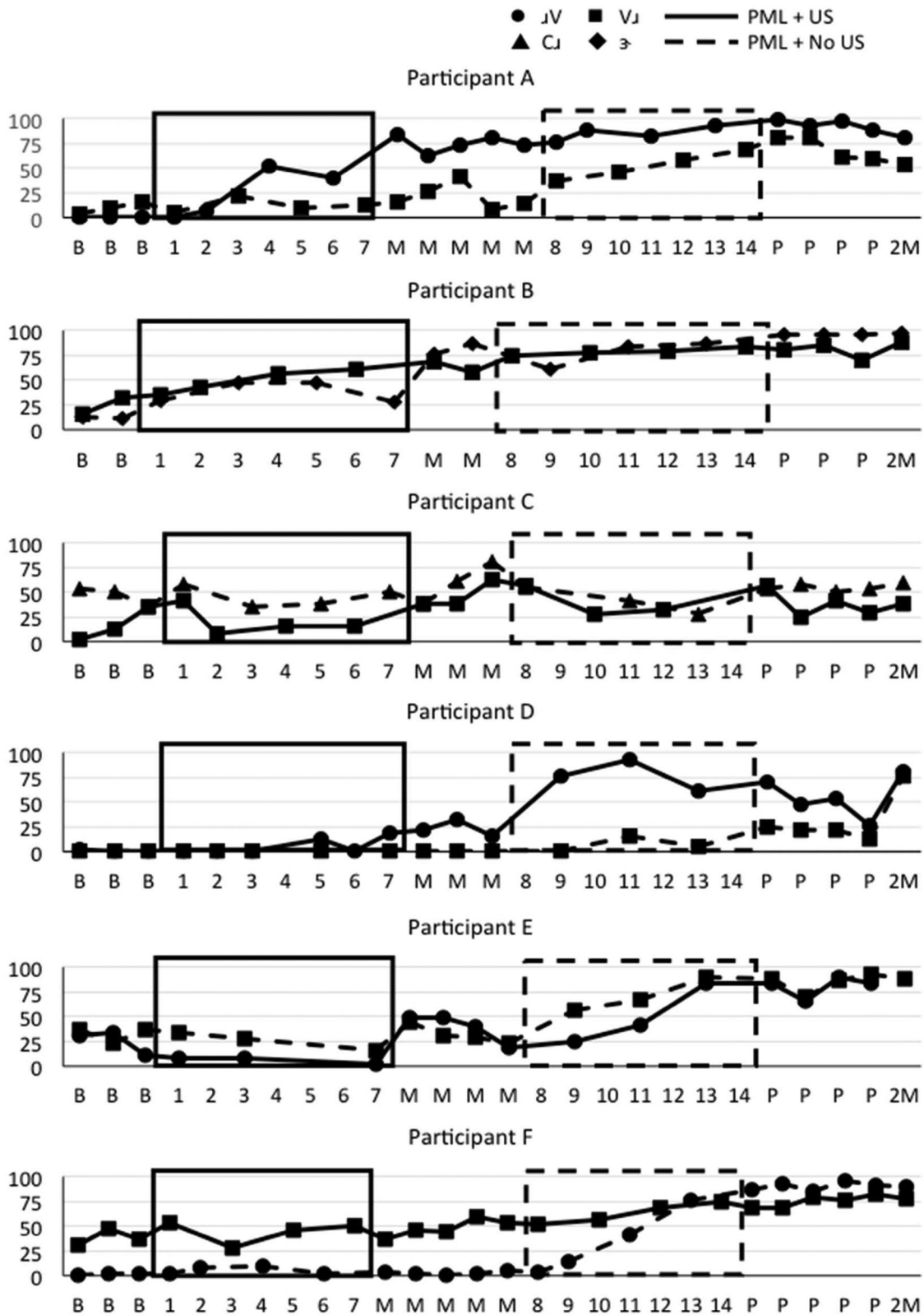


Figure 2. Generalization probe data for six participants who were treated with Principals of Motor Learning plus Ultrasound (PML+US) in the first treatment Phase and Principals of Motor Learning plus No Ultrasound (PML+NoUS) in the second Phase. Participants are ordered by treatment response in the first phase of treatment. Solid boxes represent sessions in which the PML + US treatment was delivered; dashed boxes represent sessions in which PML + NoUS treatment was delivered. Solid lines represent treatment targets treated with PML + US; dashed lines represent treatment targets treated with PML + NoUS. B = baseline (pre-treatment), M = midpoint between phases, P = post-treatment.

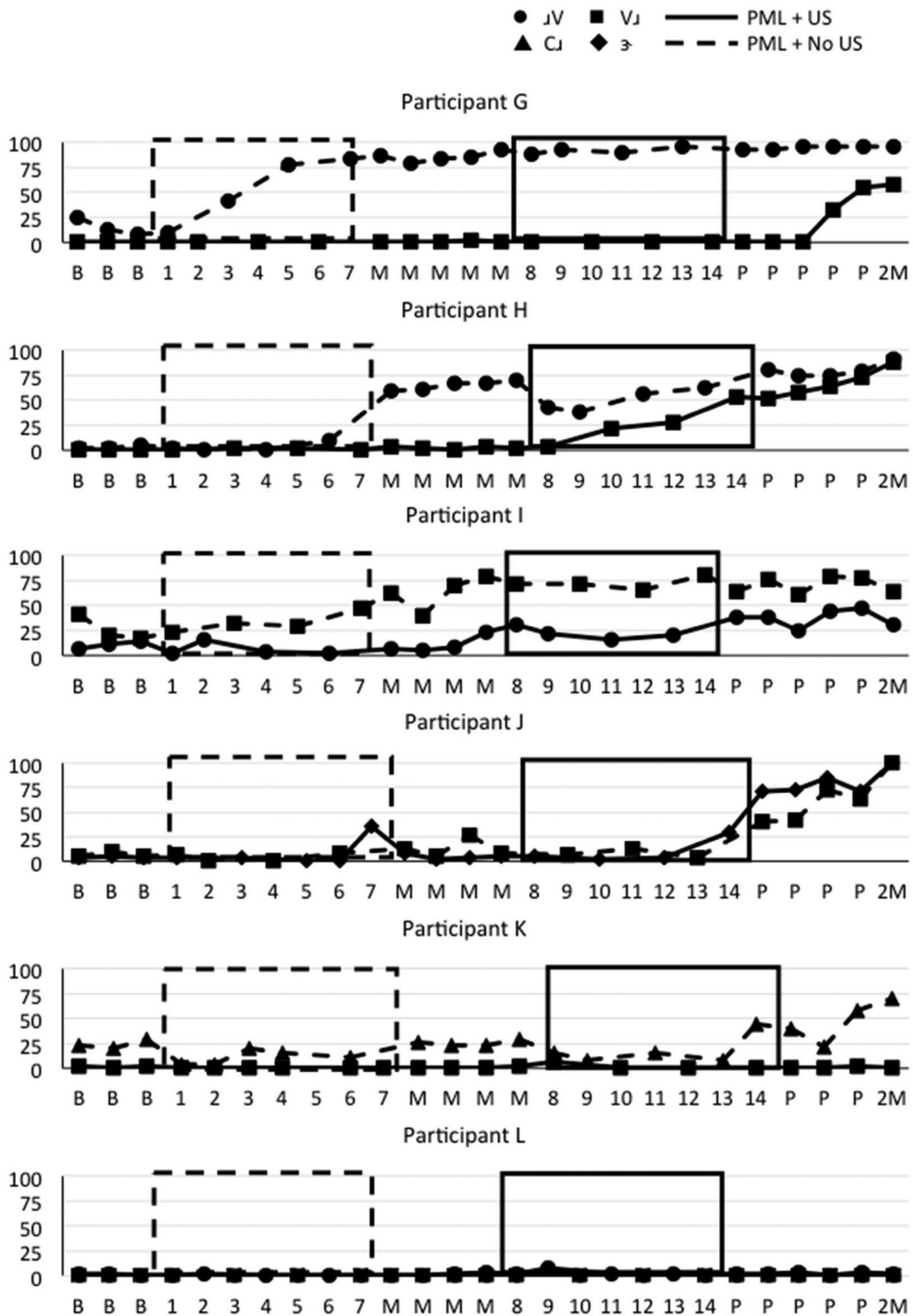


Figure 3. Generalization probe data for six participants treated with principals of motor learning plus no ultrasound (PML + NoUS) in the first treatment phase and PML plus ultrasound (PML + US) in the second phase. Participants are ordered by treatment response in the first phase of treatment. Solid boxes represent sessions in which the PML + US treatment was delivered; dashed boxes represent sessions in which PML + NoUS treatment was delivered. Solid lines represent treatment targets treated with PML + US; dashed lines represent treatment targets treated with PML + NoUS. B = baseline (pre-treatment), M = midpoint between phases, P = post-treatment.

Table 2. Treatment Response by Participant

| Participant | Phase I |                    |                |                | Phase II |                   |                  |                |
|-------------|---------|--------------------|----------------|----------------|----------|-------------------|------------------|----------------|
|             | Target  | Condition          | % increase     | d <sub>2</sub> | Target   | Condition         | % increase       | d <sub>2</sub> |
| A           | /ɪV/    | PML + US           | 74.9           | 20.59          | /Vɪ/     | PML + NoUS        | 47.0             | 3.71           |
| B           | /Vɪ/    | PML + US           | 39.4           | 4.01           | /ɜ/      | PML + NoUS        | 15.0             | 4.24           |
| C           | /Vɪ/    | PML + US           | 25.6           | 1.38           | /Cɪ/     | PML + NoUS        | -4.4             | -0.43          |
| D           | /ɪV/    | PML + US           | 21.9           | 4.79           | /Vɪ/     | PML + NoUS        | 19.7             | 6.45           |
| E           | /ɪV/    | PML + US           | 18.3           | 1.32           | /Vɪ/     | PML + NoUS        | 52.9             | 5.51           |
| F           | /Vɪ/    | PML + US           | 6.7            | 0.74           | /ɪV/     | PML + NoUS        | 87.5             | 27.24          |
| Mean (SD)   |         | Phase I US         | 31.1<br>(23.9) | 4.92<br>(7.76) |          | Phase II NoUS     | 36.3<br>(32.9)   | 7.79<br>(9.82) |
| G           | /ɪV/    | PML + NoUS         | 72.4           | 12.34          | /Vɪ/     | PML + US          | 17.1             | 1.34           |
| H           | /ɪV/    | PML + NoUS         | 58.4           | 10.09          | /Vɪ/     | PML + US          | 59.1             | 11.19          |
| I           | /Vɪ/    | PML + NoUS         | 39.2           | 3.04           | /ɪV/     | PML + US          | 24.3             | 2.44           |
| J           | /Vɪ/    | PML + NoUS         | 6.8            | 1.15           | /ɜ/      | PML + US          | 69.7             | 14.56          |
| K           | /Cɪ/    | PML + NoUS         | 4.2            | 0.53           | /Vɪ/     | PML + US          | -0.9             | -0.43          |
| L           | /Vɪ/    | PML + NoUS         | 0              | 0              | /ɪV/     | PML + US          | 0.3              | 0.15           |
| Mean (SD)   |         | Phase I PML + NoUS | 30.2<br>(31.0) | 4.52<br>(5.33) |          | Phase II PML + US | 28.25<br>(29.79) | 4.88<br>(6.37) |
| Mean (SD)   |         | Phase I total      | 30.7<br>(26.4) | 4.72<br>(6.35) |          | Phase II total    | 32.3<br>(30.2)   | 6.33<br>(8.03) |
| Mean (SD)   |         | PML + US total     | 29.7<br>(25.8) | 4.90<br>(6.77) |          | PML + NoUS total  | 33.2<br>(30.6)   | 6.16<br>(7.72) |

Note: PML = principles of motor learning; US = ultrasound; NoUS = no ultrasound.

was no consistent advantage of one treatment over another across all participants on the generalization probes (table 2).

#### *Further evidence of retention and generalization*

The 2-month follow-up data shown in figures 2 and 3 are not included in the averages from table 2, but the generalization probes reveal continued growth and retention. For example, participants D–F showed the least improvement from immediately before to immediately after phase I of treatment with PML + US, but in all cases their accuracy continued to increase during the second phase of treatment and remained high (above 75%) at the 2-month follow-up. Similarly, participants J and K were treated in the PML + NoUS condition in phase I and showed limited response initially, but their targets continued to improve during phase II and further increased by the 2-month follow-up.

The 15-item sentence imitation task that was administered prior to treatment was re-administered in the session following the completion of phase II. The results of a paired *t*-test suggested that there was a significant improvement in per cent rhotics correct from pre-treatment (mean = 16% SD = 13%) to post-treatment (mean = 45% SD = 24%,  $t(11) = 5.5$ ,  $p < .001$ ) indicating that the combined effects of both interventions facilitated improved sentence-level accuracy.

#### *Within-session practice*

With respect to the progress made within sessions in the two conditions, we explored how long it took par-

ticipants to pass the elicitation stage (i.e., 12 correct practice attempts with cueing). In the PML + US condition, 11 of 12 participants passed elicitation in their first session (participant D passed elicitation in the third session). In the PML + NoUS condition, nine of 12 participants passed elicitation in their first session (participants D and H passed elicitation in their second session; participant L failed to pass through elicitation in any of the seven sessions in the PML + NoUS condition) (figure 4).

The number of trials in structured practice was similar across the two conditions. In the PML + US condition, participants practised on average 2012 (SD = 917) attempts in seven sessions. In the PML + NoUS condition, participants practised on average 2103 (SD = 1164) trials in seven sessions. There was no significant difference in the average number of trials in structured practice per condition ( $t[11] = 0.32$ ,  $p = 0.759$ ).

## Discussion

Facilitating improved speech-sound accuracy in individuals with RSSEs can be a clinical challenge. The results of the study suggest that the two approaches used here, PML + US and PML + NoUS, could both facilitate measurable learning in only seven sessions of therapy for many children. Because the single case design was replicated among 12 participants, we had the opportunity to observe both individual response to treatment and an average increase as a result of the two treatments. The improvement observed here—approximately a 30% increase in generalization probe scores after seven sessions

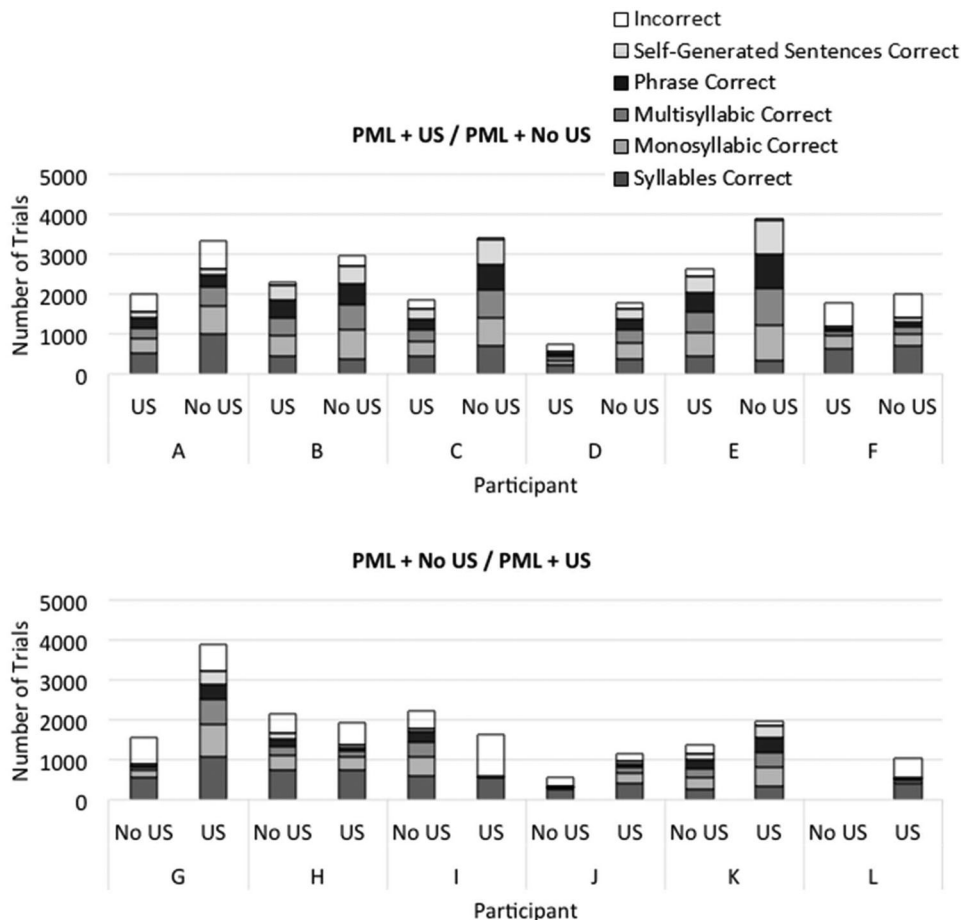


Figure 4. Number of trials scored correct by treating clinician in structured practice per treatment phase by participant. The top panel shows the six participants whose treatment order was PML + ultrasound followed by PML + no ultrasound. The bottom panel shows the six participants whose treatment order was PML + no ultrasound followed by PML + ultrasound. The number of correct trials at the syllable, monosyllabic word, multisyllabic word, phrase and sentence levels are shaded. The top of each bar represents the total number of practice trials attempted within the treatment phase (seven sessions). (Participant L did not pass the elicitation stage in the PML + NoUS condition and therefore shows no practice trials.)

(table 2)—is similar to the improvement seen in prior work using similar intervention procedures for a similar duration in children with RSSEs (Preston *et al.* 2014). Although longer treatment programmes are necessary in many cases, all but two participants (K and L) showed an increase greater than 15% accuracy on at least one of the treatment targets following seven sessions. Additionally, generalization to sentences was observed immediately after treatment, and the 2-month follow-up data (figures 2 and 3) provide strong evidence of retention and generalization for rhotics that were treated in both conditions. Thus, for most of these children with RSSEs a relatively short duration of therapy can lead to measurable gains when incorporating a number of PML.

With respect to our two hypotheses about the effect of visual KP feedback, there was evidence of individual response to treatment and no consistent pattern favouring one treatment condition emerged across all participants. One possibility is that seven sessions of

treatment is not a sufficiently long duration to observe consistent differences across treatments; longer duration of treatment in each phase might be necessary for differences to emerge. Alternatively, the similar gains with and without ultrasound visual feedback could mean that the ultrasound visual feedback was unnecessary. However, inspection of individual participant's data suggests that ultrasound visual feedback did play a role for several children, albeit with opposite effects, thus eliminating differences when considering group averages.

#### *Individual response*

The individual responses to therapy likely indicate that some children and/or treatment targets can benefit more from ultrasound feedback whereas others may be adequately treated by a PML-based intervention without ultrasound. According to hypothesis 2 (KP impedes learning), the added KP in the form of ultrasound

feedback was expected to hinder learning, perhaps due to an overreliance on such feedback or increased cognitive demands associated with processing ultrasound feedback (Hodges and Franks 2001). Several children did demonstrate poorer retention and generalization for the PML + US condition, although the specific underlying reason cannot be determined based on these findings. In contrast, several other children showed an advantage for the PML + US condition, consistent with hypothesis 1 (KP is beneficial). One possible interpretation is that for these children, the movement targets for accurate production of rhotic sounds may not have been clear, and the visual ultrasound KP may have helped to establish an internal reference of correctness for their productions and enable them to solve this complex motor problem. Overall, the fact that many children showed signs of acquisition and/or generalization with PML + US, and that some of the participants showed greater generalization with PML + US, suggests that the use of ultrasound should not be discarded based on this study. Rather, participant-specific factors and alternate research designs should be considered to more fully address relative benefit of ultrasound in speech-sound therapy.

For instance, stimulability by syllable position could play a role in rate of improvement. As an example, we qualitatively observed that participant G was readily stimulable for (i.e., able to immediately imitate successfully) onset singleton /ɪV/ (which was assigned to PML + NoUS condition) but not for coda /Vɪ/ (which was assigned to PML + US condition), and the relative advantage for PML + NoUS for this participant may be more indicative of stimulability than the treatment conditions. Hence, factors other than the use of ultrasound could influence learning and would require that future research designs take into account stimulability when assigning treatment targets and treatment conditions. One participant who did not show generalization in either condition (i.e., participant L) showed signs of acquisition of /ɪ/ only during in the PML + US condition. That is, she never passed the elicitation phase during her sessions without the ultrasound but she did show signs of acquiring /ɪ/ during treatment with the ultrasound (figure 4); the fact that she was beginning to acquire /ɪ/ with the ultrasound might indicate that, had the duration of the intervention been longer, she could have begun to generalize her accurate productions. Moreover, this participant had age-appropriate vocabulary and expressive language but relatively poor phonological processing skills (CTOPP-2 standard score 84, SAILS 86%, Nonword Repetition per cent phonemes correct 69%; table 1). It is possible that children with this profile might need treatment that is aimed at improving awareness and perception of speech sounds with a focus on detecting phonetically acceptable and unacceptable versions of the target sound (Preston *et al.* 2015); sim-

ilar profiles were observed in children with childhood apraxia of speech who failed to show signs of acquisition and/or generalization with a PML + US approach (Preston *et al.* 2016). Therefore, factors such as stimulability and phonological processing skills may interact with treatment condition.

#### *Caveats and limitations*

Given the state of the literature, a within-subject research design comparing treatment conditions was chosen. One limitation of the design is that half of the participants were treated in the PML + US for seven sessions before they were treated in the PML + NoUS condition (although different syllable positions were targeted in each phase of treatment); for these six participants, initial exposure to ultrasound visual feedback might have resulted in increasing their understanding of lingual movements for /ɪ/ that could then transfer to their practising on /ɪ/ targets without the ultrasound. Additionally, for the two participants who began the study with relatively high /ɪ/ accuracy in onset position but had low accuracy on rhotics in nucleus and coda positions (i.e., participants B and J), the improvements in nucleus and coda position occurred simultaneously. Thus, the assumption of relative independence of the treatment targets and/or the treatment conditions may not always hold true. Therefore, the within-participant design is not without its limitations, and the appropriate way to avoid these potential confounds is through a randomized controlled trial.

Moreover, the use of ultrasound is generally viewed as a tool for facilitating acquisition of a sound and likely should be withdrawn after the client is capable of successfully achieving consistently correct productions at the syllable and word level. Thus, the preferred order of intervention over a more extended treatment program for any given target would involve PML + US followed by a transition toward PML + NoUS. Further studies on the best approach for implementing and withdrawing PML + US intervention, as well as cost-benefit analyses of technologically supported interventions, would be valuable.

#### **Summary and conclusions**

There are few studies comparing different treatment approaches for children with RSSEs. This study sought to compare PML-based interventions with and without ultrasound visual feedback. There was significant variability in the response pattern both within and across children, with some children showing evidence of learning in only one condition, some showing evidence of learning in both conditions, and some failing to reveal evidence of learning with either approach. This

suggests that a number of individual factors likely play a role, and future studies should continue to explore not only comparisons between treatment approaches but also the characteristics of children who respond better to one treatment over another. The fact that 10 of 12 participants made measurable progress in only seven sessions, despite previously unsuccessful treatment, is clinically meaningful and indicates that many children with RSSEs can continue to improve their accuracy of rhotics given appropriate PML-based intervention. The 2-month follow-up sessions suggested that the combined treatment package of 14 sessions resulted in long-term improvements for most of the participants. The study therefore provides evidence that PML-based interventions, with or without ultrasound, could be viable treatment options for children with RSSEs.

### Acknowledgements

Thanks to Marika Chudacik, Olivia Flaherty, Vanessa Harwood, Kerry McNamara, Greta Sjolie, Monique Uy, Jessica Whittle and MaryKate Zoppi for their assistance with this project. This study was supported by NIH grant R03DC013152. **Declaration of interest:** The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

### References

- ADLER-BOCK, M., BERNHARDT, B., GICK, B. and BACSFALVI, P., 2007, The use of ultrasound in remediation of North American English /r/ in 2 adolescents. *American Journal of Speech–Language Pathology*, **16**, 128–139.
- ASHA, 2014, *Schools Survey Report: SLP Caseload Characteristics* (available at: <http://www.asha.org/research/memberdata/SchoolsSurvey.htm>) (accessed on 9 December 2015).
- 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**, 206–217.
- BOYCE, S. E., 2015, The articulatory phonetics of /r/ for residual speech errors. *Seminars in Speech and Language*, **36**, 257–270.
- DELATTRE, P. and FREEMAN, D. C., 1968, A dialect study of American r's by x-ray motion picture. *Linguistics*, **6**, 29–68.
- DOLLAGHAN, C. A. and CAMPBELL, T. F., 1998, Nonword repetition and child language impairment. *Journal of Speech, Language, and Hearing Research*, **41**, 1136–1146.
- DUNN, L. M. and DUNN, D. M., 2007, *Peabody Picture Vocabulary Test—Fourth Edition* (Minneapolis, MN: Pearson).
- GOLDMAN, R. and FRISTOE, M., 2000, *Goldman–Fristoe Test of Articulation—Second Edition* (Circle Pines, MN: AGS).
- HITCHCOCK, E. R. and MCALLISTER BYUN, T., 2015, Enhancing generalisation in biofeedback intervention using the challenge point framework: a case study. *Clinical Linguistics and Phonetics*, **29**, 59–75.
- HODGES, N. J. and FRANKS, I. M., 2001, Learning a coordination skill: interactive effects of instruction and feedback. *Research Quarterly for Exercise and Sport*, **72**, 132–142.
- MAAS, E., ROBIN, D. A., AUSTERMANN HULA, S. N., FREEDMAN, S. E., WULF, G., BALLARD, K. J. and SCHMIDT, R. A., 2008, Principles of motor learning in treatment of motor speech disorders. *American Journal of Speech–Language Pathology*, **17**, 277–298.
- MAGLOUGHLIN, L., 2016, Accounting for variability in North American English /ɹ/: evidence from children's articulation. *Journal of Phonetics*, **54**, 51–67.
- MCALLISTER BYUN, T. and 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**, 207–221.
- MCALLISTER BYUN, T., HITCHCOCK, E. R. and 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**, 2116–2130.
- MCGOWAN, R. S., NITTROUER, S. and MANNING, C. J., 2004, Development of /r/ in young, Midwestern, American children. *Journal of the Acoustical Society of America*, **115**, 871–884.
- NEWELL, K., CARLTON, M. and ANTONIOU, A., 1990, The interaction of criterion and feedback information in learning a drawing task. *Journal of Motor Behavior*, **22**, 536–552.
- PRESTON, J. L., BRICK, N. and LANDI, N., 2013, Ultrasound biofeedback treatment for persisting childhood apraxia of speech. *American Journal of Speech–Language Pathology*, **22**, 627–643.
- PRESTON, J. L., IRWIN, J. R. and TURCIOS, J., 2015, Perception of speech sounds in school-aged children with speech-sound disorders. *Seminars in Speech and Language*, **36**, 224–233.
- PRESTON, J. L., MAAS, E., WHITTLE, J., LEECE, M. C. and MCCABE, P. (2016), Limited acquisition and generalisation of rhotics with ultrasound visual feedback in childhood apraxia. *Clinical Linguistics & Phonetics*, **30(3-5)**, 363–381.
- PRESTON, J. L., MCCABE, P., RIVERA-CAMPOS, A., WHITTLE, J. L., LANDRY, E. and MAAS, E., 2014, Ultrasound visual feedback treatment and practice variability for residual speech-sound errors. *Journal of Speech, Language, and Hearing Research*, **57**, 2102–2115.
- RVACHEW, S., 1994, Speech perception training can facilitate sound production learning. *Journal of Speech and Hearing Research*, **37**, 347–357.
- SEMEL, E., WIIG, E. H. and SECORD, W. A., 2003, *Clinical Evaluation of Language Fundamentals—Fourth Edition* (San Antonio, TX: Harcourt Assessment).
- SHRIBERG, L. D., 2009, Childhood speech-sound disorders: from postbehaviorism to the postgenomic era. In R. Paul and P. Flipsen (eds), *Speech Sound Disorders in Children* (San Diego, CA: Plural), 1–33.
- THOONEN, G., MAASSEN, B., GABREËLS, F. and SCHREUDER, R., 1999, Validity of maximum performance tasks to diagnose motor speech disorders in children. *Clinical Linguistics and Phonetics*, **13**, 1–23.
- WAGNER, R. K., TORGESEN, J. K., RASHOTTE, C. A. and 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).
- WILLIAMS, A. L., 2012, Intensity in phonological intervention: is there a prescribed amount? *International Journal of Speech–Language Pathology*, **14**, 456–461.
- WILLIAMS, K. T., 2007, *Expressive Vocabulary Test—Second Edition* (Minneapolis, MN: Pearson).