Clinical Linguistics & Phonetics, 2014; Early Online: 1–17

ISSN: 0269-9206 print / 1464-5076 online
DOI: 10.3109/02699206.2014.956232

Enhancing generalisation in biofeedback intervention using the challenge point framework: A case study

ELAINE R. HITCHCOCK[1] & TARA McALLISTER BYUN[2]

[1]*Department of Communication Sciences and Disorders, Montclair State University, Bloomfield, NJ, USA and* [2]*Communicative Sciences and Disorders, New York University, New York, NY, USA*

(*Received 31 May 2014; revised 13 August 2014; accepted 16 August 2014*)

Abstract
Biofeedback intervention can help children achieve correct production of a treatment-resistant error sound, but generalisation is often limited. This case study suggests that generalisation can be enhanced when biofeedback intervention is structured in accordance with a ''challenge point'' framework for speech-motor learning. The participant was an 11-year-old with residual /r/ misarticulation who had previously attained correct /r/ production through a structured course of ultrasound biofeedback treatment but did not generalise these gains beyond the word level. Treatment difficulty was adjusted in an adaptive manner following predetermined criteria for advancing, maintaining, or moving back a level in a multidimensional hierarchy of functional task complexity. The participant achieved and maintained virtually 100% accuracy in producing /r/ at both word and sentence levels. These preliminary results support the efficacy of a semi-structured implementation of the challenge point framework as a means of achieving generalisation and maintenance of treatment gains.

Keywords: Biofeedback, challenge-point framework, generalisation, principles of motor learning, speech sound disorder

Introduction

Speech sound disorder affects an estimated 10% of preschool and school-aged children (NIDCD, 1994). Unintelligible or perceptually atypical speech can pose a barrier to academic and social participation, with potentially lifelong consequences for educational and occupational outcomes (McCormack, McLeod, McAllister, & Harrison, 2009). Most developmental speech errors resolve by the time children are 8–9 years of age. When atypical speech patterns persist beyond this time frame, they are termed residual speech errors (Shriberg, 1994). These persistent errors pose a particular challenge for speech-language pathologists, with the result that many clients are discharged with residual errors uncorrected (Ruscello, 1995). Recent research has provided increasingly strong evidence that visual biofeedback intervention can succeed in eliciting correct production from children whose speech errors have not responded to other forms of treatment (Adler-Bock, Bernhardt, Gick, & Bacsfalvi, 2007; McAllister Byun & Hitchcock, 2012;

Correspondence: Elaine R. Hitchcock, Department of Communication Sciences and Disorders, Montclair State University, 1515 Broad Street, 2nd floor, Bloomfield, NJ 07003, USA. Tel: 973-229-3797. Fax: 973-655-7372. E-mail: hitchcocke@mail.montclair.edu

Modha, Bernhardt, Church, & Bacsfalvi, 2008; Preston, Brick, & Landi, 2013; Ruscello, 1995; Shuster, Ruscello, & Smith, 1992; Shuster, Ruscello, & Toth, 1995). However, many of these studies have identified a limitation of biofeedback treatment: generalisation of gains made in treatment is not automatic, with some participants remaining largely dependent on the continued availability of visual feedback to achieve correct production of their speech sound targets (Fletcher, Dagenais, & Critz-Crosby, 1991; Gibbon & Paterson, 2006; McAllister Byun & Hitchcock, 2012; McAllister Byun, Hitchcock, & Swartz, in press). Research in the challenge point framework (Guadagnoli & Lee, 2004) suggests that generalisation of motor learning can be enhanced by structuring treatment so that the learner is challenged to precisely the right degree. Subsequently, Rvachew and Brosseau-Lapré (2012) discussed the application of the challenge point framework in intervention for speech sound errors. Their findings suggest that manipulation of the conditions of practice could allow for greater generalisation of gains made through biofeedback treatment. This case study describes an 11-year-old child with residual /r/ misarticulation who attained correct production of /r/ through a structured course of biofeedback treatment but did not generalise these gains beyond the single word level. We explored the hypothesis that follow-up treatment structured in accordance with the principles of the challenge point framework would allow this participant to generalise correct /r/ production to the sentence level and to maintain these gains over time.

Biofeedback intervention

Visual biofeedback uses instrumentation to provide real-time information about aspects of speech that the speaker has limited insight into under ordinary circumstances, with the goal of bringing these processes under conscious control (Volin, 1998). In biofeedback intervention, the client's attention is directed towards a real-time visual representation of a speech sound. With guidance from the clinician, the client attempts to produce the target sound while watching the display. Using both cues from the clinician (e.g. models of the target sound or verbal articulator placement cues) and input from the visual speech sound representation, the client is taught to modify his/her own output in an effort to achieve a closer match with the visual model. Biofeedback can be provided with a range of technologies, such as electropalatography (EPG; e.g. Fletcher et al., 1991) or real-time spectral or spectrographic displays (e.g. McAllister Byun & Hitchcock, 2012; Shuster et al., 1992, 1995). The present study used ultrasound imaging as the mechanism of biofeedback delivery. In ultrasound biofeedback, an ultrasound probe positioned beneath the client's chin allows the client to view the shape and movements of his/her tongue, which he/she attempts to modify for a closer match with a visual model representing correct articulation of the target sound.

Previous studies of ultrasound biofeedback treatment for residual errors affecting /r/, while somewhat limited in number and scope, suggest that this method holds a great deal of promise.[1] Adler-Bock et al. (2007) found that two adolescents with persistent /r/ misarticulation made substantial gains in perceptually rated accuracy of /r/ at the word level after 14 one-hour sessions of ultrasound biofeedback treatment. Modha et al. (2008) reported similar results in a case study of one 13-year-old male with persistent /r/ misarticulation who received a combination of ultrasound biofeedback and traditional articulatory treatment. Finally, Preston et al. (2013) conducted a multiple-baseline study investigating the effects of ultrasound biofeedback on the

[1]Omitted in the interest of brevity are several investigations documenting the efficacy of ultrasound biofeedback intervention for different populations, including individuals with hearing impairment (Bacsfalvi, 2010; Bacsfalvi, Bernhardt, & Gick, 2007; Bernhardt, Gick, Bacsfalvi, & Ashdown, 2003; Shawker & Sonies, 1985) and Down Syndrome (Fawcett, Bacsfalvi, & Bernhardt, 2008).

production of /r/ and other phonemes in individuals with childhood apraxia of speech (CAS). Preston et al. reported that in five of six participants, production of /r/ improved significantly following the introduction of ultrasound biofeedback intervention.

Generalisation in biofeedback intervention

There is increasingly robust evidence that biofeedback can succeed in eliciting correct /r/ from children whose /r/ misarticulation has not responded to other forms of intervention. However, some biofeedback treatment studies have noted a limited degree of spontaneous generalisation of skills acquired through biofeedback treatment to a context in which biofeedback is not available (Fletcher et al., 1991; Gibbon & Paterson, 2006; McAllister Byun & Hitchcock, 2012; McAllister Byun et al, in press). Gibbon and Paterson (2006) conducted a survey in which speech-language pathologists were asked to describe the outcomes achieved by 60 children who had received EPG biofeedback intervention for speech deficits with varying aetiologies. The survey results indicated that 87% of participants had made progress in their speech accuracy over the course of biofeedback treatment; however, 88% of participants were characterised as demonstrating at least some degree of difficulty in generalising their newly acquired articulatory skills into the context of spontaneous speech.

Studies using EPG biofeedback for speech sound errors have reported that biofeedback is most effective in early stages of therapy, when a target sound is first being established (Fletcher et al., 1991; Gibbon & Paterson, 2006). The notion that biofeedback is more effective for establishing than for generalising motor patterns receives further support from the broader body of research investigating general principles of motor skill learning. Certain conditions of practice and feedback have proved more facilitative of the initial acquisition of a motor plan, while different conditions may maximise retention and transfer (see review in Maas et al., 2008). Most notably, qualitative knowledge of performance feedback – of which biofeedback is one subtype – has been found to be effective when the motor task is novel or the nature of the target is unclear to the learner (Newell, Carlton, & Antoniou, 1990). However, in later stages of learning, knowledge of performance feedback may lose its advantage or even become detrimental to learning (Maas et al., 2008).

The challenge point framework

Recent research in the challenge point framework (Rvachew & Brosseau-Lapré, 2012) suggests that manipulation of the conditions of practice could allow for greater generalisation of gains made through biofeedback treatment. The challenge point concept is derived from motor learning research by Guadagnoli and Lee (2004), who define the challenge point as the moment when ''the factors contributing to functional task difficulty (including the level of the performer and practice conditions) interact to dictate the optimal amount of interpretable information and, hence, the potential for learning'' (p. 216). In some models of motor learning, errors provide the motor control system with an opportunity to update its parameter settings; they are thus a vital part of the learning process. If functional task difficulty is too low and accuracy during practice is very high, the child does not receive any error feedback, and opportunities for learning are not created. On the other hand, if functional task difficulty is too high and accuracy is very low, the learner may be overwhelmed by complexity and is not likely to be able to encode information that can be used in future motor performance. Thus, learning will be maximised when the level of difficulty is neither too high nor too low. The challenge point concept is related to the notion of the ''zone of proximal development'' in cognitive learning (Vygotsky, 1978).

Rvachew and Brosseau-Lapré (2012) have adapted the challenge point concept to the specific context of speech development and disorders. Adopting a dynamic systems model of speech development, Rvachew and Brosseau-Lapré identify numerous factors that contribute to the functional difficulty of a speech performance task, including the child's skill level, difficulty of the target, level of clinician support, practice schedule, and the nature of feedback provided. In our previous biofeedback studies, most of these parameters were held fixed at a low or easy level throughout the treatment period. For example, practice was blocked, targets were modeled by the clinician, and informative feedback was provided with high frequency. From the point of view of the challenge point framework, then, it is unsurprising that this treatment produced a high level of within-practice accuracy but afforded little generalisation to other contexts.

The present study was thus designed to explore the hypothesis that the application of challenge point principles during a period of biofeedback intervention may enhance generalisation for a child with residual speech errors affecting /r/. The study had a secondary goal of evaluating the feasibility of a semi-structured implementation of the challenge point framework. Providing treatment within the challenge point framework can be daunting for anyone other than an experienced clinician: a learner's optimum challenge point can shift from trial to trial within a treatment session, and the clinician must constantly adjust to maintain the appropriate level of difficulty. Treatment in the present study followed a standard flowchart, with predetermined levels of accuracy for advancing, maintaining, or moving back a level in a multidimensional hierarchy of functional task complexity. If this method proves to be effective, it could increase the impact of challenge point treatment by making it accessible to clinicians with a broader range of experience levels.

Residual errors affecting /r/

This study examines residual errors affecting the phoneme /r/, a late-emerging sound that is one of the most common treatment-resistant errors in English (Ruscello, 1995; Smit, Hand, Freilinger, Bernthal, & Bird, 1990). The difficulty children experience in acquiring /r/ is presumed to be at least partially attributable to the complexity of the articulatory configuration for /r/ (Gick et al., 2008). While most speech sounds are produced with only a single major narrowing of the vocal tract, English /r/ is produced with two major lingual constrictions: one in which the anterior tongue approximates a point near the hard palate, and one in which the tongue root retracts to narrow the pharyngeal space. Many speakers additionally round the lips (e.g. Adler-Bock et al., 2007) or form a midline groove with the tongue (e.g. Bacsfalvi, 2010), further increasing the complexity of /r/ articulation.

English /r/ can appear as a consonant in syllable onset position in words like *red* and *green*. For simplicity, we follow the convention of using /r/ rather than /ɹ/ to transcribe the English consonantal rhotic. It can also act as the nucleus of a syllable, as in /hɝ/ (*her*) or /fivɚ/ (*fever*). Finally, /r/ can occur in postvocalic position in words like *hair* and *near*. McGowan, Nittrouer, and Manning (2004) characterised this variant as the vocalic offglide of a rhotic diphthong (e.g. /hɛɚ/, /nɪɚ/), offering evidence that post-vocalic /r/ is acoustically and articulatorily more similar to syllabic /r/ than onset /r/. Furthermore, studies of intervention for /r/ frequently report generalisation between syllabic /r/ and rhotic diphthongs, whereas spontaneous generalisation between these variants and onset /r/ tends to be more limited (McAllister Byun & Hitchcock, 2012; Preston et al., 2013). Based on this evidence, the present article will group syllabic and post-vocalic /r/ under the broad category ''vocalic /r/,'' with the term ''consonantal /r/'' applied to /r/ in syllable onset position only.

Case study

Participant

The subject of this case study, pseudonym ''Lilianne,'' is a native monolingual English-speaking female who was aged 11;2 at the start of the present study. Lilianne was identified by referral from a local speech-language pathologist in response to community postings describing our previous study of the efficacy of ultrasound biofeedback intervention (McAllister Byun, Hitchcock, & Swartz, in press). At the time of her initial referral, Lilianne was aged 10; 9 and had received 2.5 years of individual speech treatment for /r/, /s/, and /z/ errors using a traditional articulatory method of intervention. Based on a parent questionnaire, Lilianne's general health and birth history were unremarkable. For inclusion in the previous study, Lilianne was required to pass a pure-tone hearing screening (1000, 2000, 4000 Hz at 20 dB HL), demonstrate no gross structural or functional abnormality in the oral mechanism, demonstrate at least 95% accuracy on the Percent Consonants Correct-Revised measure (PCC-R; Shriberg, Austin, Lewis, McSweeny, & Wilson, 1997) after /r/ targets were excluded, and score below 30% accuracy on a standard stimulability probe for /r/ (Miccio, 2002). To establish that receptive language was adequate to understand study directions, Lilianne was required to score within one standard deviation of the age-level mean on the ''Auditory Comprehension'' subtest of the *Test of Auditory Processing Skills-3* (Martin & Brownell, 2005). Lilianne received a standard score of 11, placing her in the 63rd percentile for her age. Finally, Lilianne was required to score below 30% accuracy on a 64-item single-word /r/ probe task, described in detail below.

Previous treatment

As a participant in the treatment study reported by McAllister Byun, Hitchcock, & Swartz (in press), Lilianne attended two 30-min ultrasound biofeedback treatment sessions per week for 8 weeks. Treatment was administered in individual sessions by a certified speech-language pathologist. In these sessions, participants practised /r/ at the syllable level while viewing an ultrasound image generated with an Interson SeeMore USB-powered ultrasound probe. Each session elicited 30 trials of isolated syllabic /r/ and 10 trials of each of the syllables /rɑ/, /ri/, and /ru/. Stimuli were elicited in constant order in blocks of five trials. Participants who demonstrated a high level of accuracy within the treatment setting were eligible for advancement through a hierarchy in which /r/ targets were embedded in non-word syllables and words of increasing complexity. Stimuli continued to be presented in a blocked fashion, with constant order of elicitation.

In a pre-treatment baseline period, Lilianne demonstrated less than 10% accuracy at the word level across both vocalic and consonantal /r/ targets elicited without feedback. After a few false starts, she responded very well to ultrasound biofeedback treatment. By her final treatment session, the treating clinician's within-session scoring indicated that Lilianne had achieved 93% accuracy for syllabic /r/ in CVC non-words and 70% accuracy for onset /r/ in CV syllables. Maintenance probes elicited without biofeedback and scored by blinded listeners confirmed a high level of accuracy for vocalic variants of /r/ in untreated words, with a mean of 96% correct across three maintenance sessions (Figure 1). This resulted in an unstandardised effect size (mean level difference) of 94.9 percentage points and a very large standardised effect size (d_2; see discussion below) of 16.7. However, Lilianne made less robust gains for consonantal /r/ targets at the word level. Her mean accuracy of 39% correct across three maintenance sessions yielded an unstandardised effect size of 34 percentage points and a standardised effect size of 4.0. Finally, a conversational sample was collected after the end of Lilianne's initial treatment period and scored for /r/ production accuracy by a speech-language pathologist who was not familiar with the study

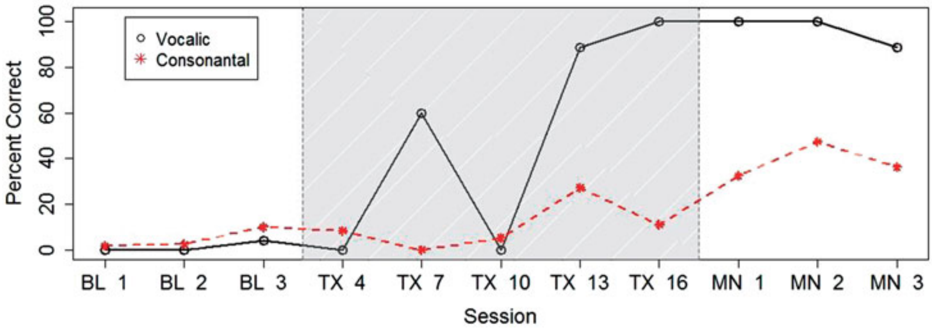

Figure 1. Lilianne's accuracy in producing untreated /r/ words elicited without feedback in a previous biofeedback treatment study (McAllister Byun, Hitchcock, & Swartz, in press). X-axis represents probe elicitation session (BL = baseline, TX = within-treatment, MN = maintenance). Y-axis represents per cent of tokens rated perceptually correct based on mode across three blinded clinician listeners.

or the participant. This clinician's scoring indicated that Lilianne produced /r/ sounds in conversation with 46% accuracy (24 of 52 /r/ targets correct), suggesting a limited degree of generalisation to connected speech for both vocalic and consonantal /r/. Thus, arrangements were made for Lilianne to return for a second course of treatment, reported in the present study. Due to scheduling conflicts, a period of 8 weeks separated the first and second treatment programmes.

Study design

The present study utilised a 17-week single case study design. Lilianne completed three baseline sessions over a period of 3 weeks prior to the onset of treatment. Each baseline session featured a single-word /r/ probe measure in which pictures and written words were used to elicit 64 familiar words containing consonantal and vocalic /r/ in various phonetic contexts. Vocalic /r/ was probed with four items representing each of the following variants: (1) stressed /ɝ/, (2) unstressed /ɚ/, (3) /ɑr/, (4) /ɛr/, (5) /ɔr/, (6) /ɪr/. Consonantal /r/ variants, which were elicited in front and back vowel contexts in equal numbers, included singleton /r/ and clusters featuring alveolar (/tr, dr/), velar (/kr, gr/), and labial place (/pr, br, fr/). No feedback was provided during probe administration. Within the treatment phase of the study, the 64-word probe was scheduled to be re-administered every fourth session, or at any point when the participant was eligible for advancement to a new target based on accuracy within practice. After the end of the treatment period, the probe was re-administered in three maintenance sessions over a period of 3 weeks and at one 1-month follow-up session. Finally, a probe eliciting /r/ at the sentence level (Schmidlin & Boyce, 2010, unpublished) was administered in the first baseline session, all within-treatment sessions, the final maintenance session, and the 1-month follow-up session. This probe consisted of five sentences containing both vocalic and consonantal /r/ in various phonetic contexts, with multiple /r/ words per sentence.

Treatment was provided in weekly individual sessions of 30–45 min in duration. Sessions elicited 60 trials of /r/ in selected phonetic contexts at the word, phrase, or sentence level. Practice within a session targeted at least three variants of /r/, chosen based on probe measure accuracy following a method discussed in more detail below. Prior to the start of the session, two words were randomly selected to represent each target, for a total of six words that would typically be elicited 10 times each. However, new target words could be introduced within a session if the participant moved to a higher level of target complexity. Each session began with a period of free play in which Lilianne could practise any /r/ target while viewing the ultrasound biofeedback

Table 1. Comparison of parameters across previous biofeedback treatment study (McAllister Byun, Hitchcock, & Swartz, in press) and current treatment study using challenge point framework.

Context	Original biofeedback study	Challenge point study
Pre-treatment assessment	• 3 baseline sessions	• 3 baseline sessions
Treatment duration	• 3 instructional sessions (45 min. each) • 16 thirty-min practice sessions over 8.5 weeks • 2 sessions per week	• 0 instructional sessions • 11 thirty- to 45-min practice sessions over 11 weeks • 1 session per week
Treatment sessions	• 60 trials elicited • Individualised cues during pre-practice and treatment trials • Initial targets: 30 trials /ɝ/, 30 trials /rɑ, ri, ru/ • Target complexity increased after >80% accuracy was achieved for consonantal or vocalic targets, summed over 30 trials; no alteration otherwise	• 60 trials elicited • Individualised cues during pre-practice and treatment trials • Initial targets: Based on most facilitative context, moving to less facilitative then least facilitative as progress was documented. • Target complexity increased, remained constant, or decreased if participant achieved >80%, 51–79% or <50% accuracy, respectively, summed over 30 trials.
Within-treatment assessment	• Every third session until end of treatment period	• Every fourth session and/or when participant reached the top of the treatment hierarchy, until discharge criterion met (>90% accuracy sustained over 2 within-treatment probes)
Post-treatment assessment	• 3 maintenance sessions immediately post-treatment	• 3 maintenance sessions immediately post-treatment; one 1-month follow-up session

display. The free play period was used as an opportunity to try out new targets (word shapes, prosodic patterns, etc.) that were likely to be encountered within a session. After 5 min of free play, /r/ targets were elicited in blocks of five trials. At the start of each block, the clinician provided verbal cues focusing the participant's attention on some aspect of the speech target. Cues typically described one component of correct articulator placement for /r/ (e.g. ''Focus on moving the back part of your tongue back''). After each block, the clinician offered feedback in the form of a general comment on the perceptual accuracy of the participant's output. Brief breaks were offered after every 10 trials. Table 1 offers a side-by-side comparison of the parameters of treatment as implemented in the original biofeedback study (McAllister Byun, Hitchcock, & Swartz, in press) versus in the present study.

Challenge point treatment

Treatment was structured according to a challenge point hierarchy in which predetermined accuracy criteria were applied to advance, maintain, or move back a level in a multidimensional hierarchy of functional task complexity. Three parameters were adjusted on a rotating schedule based on accuracy *within* a treatment session, and one additional parameter could be adjusted *across* sessions. These parameters are described in detail below. Production accuracy within sessions was determined by the clinician administering treatment (either the first author or another trained speech-language pathologist). The treating clinician entered her scores in a standard Excel spreadsheet (''session spreadsheet'').

Lilianne's initial session began with all parameters at the lowest possible level of the hierarchy. The starting point in subsequent sessions was based on her performance in the previous session.

Table 2. Parameters making up the multidimensional hierarchy of the present implementation of the challenge point framework, adjusted on a rotating basis.

A. Feedback frequency	B. Mode of elicitation	C. Target complexity	D. Order of elicitation
I. 100%	I. Imitate clinician's model	Ia. 1 syllable, simple Ib. 1 syllable, complex (competing /l/ or /w/)	I. Fully blocked
II. 50% (one block on, one block off)	II. Read independently (no prosodic manipulation)	IIa. 2 syllables, simple IIb. 2 syllables, complex (competing /l/ or /w/)	II. Random-blocked
III. 0%	III. Imitate clinician's model with prosodic manipulation	IIIa. 3 syllables, simple IIIb. 3 syllables, complex (competing /l/ or /w/)	III. Fully random
	IV. Read independently with prosodic manipulation	IV. 1-, 2- or 3-syllable words (simple or complex) in carrier phrase V. 1-, 2- or 3-syllable words (simple or complex) in sentences VI. Sentences with multiple /r/ targets	

The parameter adjusted across rather than within sessions is shaded grey.

Within treatment sessions, the session spreadsheet automatically tallied the number of items scored correct in each block of trials. After every two blocks of five trials, the per cent of correct responses across those previous two blocks (10 trials) was calculated and used to make a determination regarding movement in the challenge point hierarchy. If accuracy in the previous 10 trials was 80% or higher, one within-treatment parameter was adjusted to a higher level of complexity. If accuracy across 10 trials was 50% or lower, one within-treatment parameter was adjusted to a lower level of complexity. For accuracy values from 51–79%, the current level of difficulty was maintained. The three within-session parameters were organised in columns, as shown in Table 2. These parameters were adjusted on a rotating basis, such that the participant's first increase in complexity affected the parameter in Column A; the next increase affected the parameter in Column B; the next in Column C; and the next increase rotated back to Column A. Whenever a decrease in complexity was called for, this change would undo the most recent adjustment. For example, if the participant scored above 80% in a block of 10, leading to an increase in complexity in Column A, but then dropped to less than 50% accuracy, the next modification would reverse the most recent change by reducing complexity in Column A.

The within-session manipulations of complexity fell into three categories: frequency with which ultrasound biofeedback was provided, mode of elicitation, and word shape/elicitation context. The parameter of ultrasound feedback frequency had three levels: 100%, 50%, and 0% biofeedback. At the 50% level, withdrawal of ultrasound feedback was scheduled such that biofeedback was provided in one block of five trials and not in the next block.

The ''mode of elicitation'' parameter had four levels. At the most facilitative level, the participant was cued to imitate the clinician's model of the target utterance. At the next level, the participant read target utterances independently, followed by a level in which the participant imitated the clinician's verbal model of a production with modified prosody. Prosodic manipulations included interrogative intonation and loud vocal volume, as well as normal production. At the most complex level, the participant was instructed to read target utterances independently using modified prosody, which was cued with written prompts on index cards. Prosodic manipulations were randomly assigned to each utterance within a block of five.

The ''target complexity'' parameter had a total of six levels. The most basic levels elicited the /r/ target in individual words, starting with simple monosyllables, then monosyllabic words featuring a ''competing'' phoneme (/l/ or /w/, which share articulatory gestures with /r/ and thus can potentially interfere with /r/ articulation). The next levels were simple disyllables and disyllables with a competing phoneme, followed by simple trisyllables and trisyllables with a competing phoneme. The next level elicited monosyllabic words (with or without a competing phoneme) in a carrier phrase, ''Say [target] again,'' followed by a level in which disyllabic and then trisyllabic words were elicited in the same carrier phrase. At the next highest level, target words were elicited in a set of five standard sentences (e.g. ''[target] is the thing that she told me to say.'') At the highest level, these sentences were minimally altered so that multiple /r/ words were produced in a single sentence (e.g. ''[target] is the word that she told me to say.'').

One additional parameter, order of target elicitation, was adjusted based on the participant's cumulative accuracy across an entire session. If the participant's accuracy across all trials in the previous session was 80% or higher, the level of complexity of this parameter was increased; if cumulative accuracy was below 50%, complexity was decreased; otherwise, this parameter was held unchanged. The lowest level of the hierarchy was a fully blocked order in which each target word was presented in two consecutive blocks of five trials, and words representing the same /r/ variant (e.g. *barn, hard* for the rhotic diphthong /ɑr/; *sir, shirt* for /ɝ/) were elicited in succession. The next level of complexity featured random-blocked order of elicitation, in which each block elicited five trials of a single word, but across blocks, different words and variants could be presented in random order. At the highest level of complexity, fully random order, different words and /r/ variants could be represented within a single block of five trials.

Target selection and discharge criteria

Treatment targets were selected based on accuracy in the 64-word /r/ probe elicited in baseline sessions. The authors independently listened to the baseline probes and then reached a consensus regarding the relative accuracy of different variants. Like the parameters manipulated within treatment sessions, treatment targets were selected to follow a hierarchy starting with the most facilitative and progressing to the least facilitative contexts. Possible variants considered were the rhotic diphthongs, stressed syllabic /ɝ/, singleton /r/ before non-back vowels, and singleton /r/ before back vowels.

Advancement to a new /r/ target within treatment was based on a combined time and performance criterion. The 64-word /r/ probe and sentence probe were administered at the start of every fourth session, and any /r/ targets that had attained a level of accuracy equivalent to the current treated targets would be added to the set of variants targeted in treatment. In addition, probe re-administration was scheduled to occur whenever the participant reached the highest level of the hierarchy for all within- and across-session parameters for her current targets. To be discharged from treatment, the participant was required to demonstrate at least 90% accuracy across all targets in both word- and sentence-level probe measures.

Measurement

All target words were isolated from audio recordings of word probes elicited in baseline, within-treatment, maintenance, and follow-up probe sessions. The full set of items ($n = 576$) was pooled, randomised, and subdivided into three blocks of approximately 200 items. Three raters, certified speech-language pathologists who were not familiar with the experiment or the participant, completed all three blocks in a self-paced fashion over the span of one or more weeks.

Word stimuli were presented in a blinded, randomised fashion using the online experiment presentation platform Experigen (Becker & Levine, 2010). Raters heard a sound file and saw the target word in conventional orthography. They could listen to each target up to three times. They then entered their accuracy rating by clicking ‘‘correct’’ or ‘‘incorrect.’’ Each item was rated by all three listeners, and the three ratings were reduced to a single accuracy score reflecting the mode across listeners. Raters were instructed to follow a strict standard whereby even distorted sounds with some rhotic quality were rated incorrect. To encourage a uniform rating standard, an initial training block was provided in which listeners rated 20 sample items and received immediate feedback based on the rating assigned by an experienced clinician in a previous study. After the training block, listeners completed an eligibility test consisting of 100 items that had been rated by the aforementioned experienced clinician. Only listeners who demonstrated $\geq$80% agreement with the experienced clinician’s ratings were retained as raters. Pairwise inter-rater reliability was calculated to be 88% for raters 1 and 2, 89% for raters 1 and 3, and 86% for raters 2 and 3.

The accuracy of /r/ words elicited with the sentence probe (Schmidlin & Boyce, 2010, unpublished) was also evaluated using the same blinded listeners’ ratings in Experigen. Because sentences contained multiple /r/ targets, each sentence was presented multiple times, with a single word orthographically displayed on the screen so the listener would know which word to focus on. Raters could listen to a sentence up to three times for each target word being evaluated. Again, the three raters’ scores for each word were reduced to a single modal value.

Data analysis

For visual inspection of changes over the course of treatment, the percentage of items rated correct was plotted across baseline, within-treatment, maintenance, and follow-up probes for both word- and sentence-level stimuli. In addition to visual inspection, standardised and unstandardised effect sizes were calculated for both vocalic and consonantal /r/ variants at the word level. Standardised effect sizes were calculated using Busk and Serlin’s (1992) d_2 statistic, reported in (1) below.

$$\text{Effect size} = \frac{\text{mean}_{\text{MN}} - \text{mean}_{\text{BL}}}{\text{SD}_{\text{pooled}}} \tag{1}$$

As described by Beeson and Robey (2006), by pooling standard deviations across baseline and maintenance periods, this measure reduces the number of cases where effect size cannot be calculated due to zero variance in the baseline phase. In cases where variance is very low, standardised effect sizes may overestimate the magnitude of the effect. Therefore, unstandardised effect sizes (raw differences between the mean percentage of items rated correct in maintenance versus baseline intervals) were also calculated.

It is important to note that although we report effect sizes indicating the magnitude of change from pre- to post-treatment, the design of the present study does not make it possible to conclude that these changes were caused by the treatment provided. That is, because experimental control was not established in this case study context, we cannot rule out the possibility that the observed changes could be attributed to external influences such as maturation. This issue is taken up in greater detail in the Discussion section.

Fidelity

Fidelity in implementation of the specified treatment protocol was measured in 18% of all sessions (Kaderavek & Justice, 2010). To measure fidelity, the audio recording of selected treatment

Table 3. Results of fidelity check.

Block consisted of five trials (%)	Interruptions during a block (%)	Qualitative feedback presented after block (%)
95.8%	4.1%*/ 37.5%	91.6%

*This value excludes cases when the participant was asked to repeat an utterance to reflect the requested prosodic change.

sessions was reviewed by research assistants not involved in treatment delivery. These raters completed a checklist to verify the following aspects of study design: (1) each block consisted of precisely five trials, (2) feedback or other interruptions did not occur within a block, and (3) feedback in the form of a general comment on the perceptual accuracy of the target production was provided after each block.

Results of the fidelity check are reported in Table 3. The primary deviations from the stated protocol involved within-block interruptions, which were observed in nearly 40% of blocks. Further inspection revealed that the great majority of these were cases where the participant initially failed to apply the prosodic manipulation indicated for a particular trial and was prompted to repeat in the correct prosodic condition. When these cases were excluded, interruptions were found in only 4% of blocks.

Results

Progress within treatment

The /r/ variants selected for Lilianne's initial phase of treatment, representing maximally facilitative contexts, were /ɝ/, /ɑr/, and /ɔr/. These targets were already produced with a high level of accuracy at the word level, and Lilianne moved rapidly through higher levels of complexity. By the end of her fourth treatment session, she had reached the top of the complexity hierarchy for all within- and across-session parameters. Thus, treatment for these targets was discontinued and three /r/ variants representing a less facilitative context (/ri/, /reɪ/, /raɪ/) were selected to be targeted in subsequent treatment sessions. All other parameters of the challenge point framework were reset to the beginning of the hierarchy.

Lilianne's progress was reassessed after treatment session 8, in accordance with the time-based criterion. At this point, she had reached the top of the complexity hierarchy for biofeedback frequency and mode of elicitation, but she had not yet reached the phrase level of target complexity. However, when the word and sentence probe were re-administered, Lilianne was judged to produce consonantal /r/ before back vowels with accuracy equivalent to that seen for her current targets. Therefore, consonantal /r/ in a back vowel context (/ra/, /ro/, /ru/) was added to the set of targets for practice within treatment.

After three additional sessions, Lilianne reached the top of the complexity hierarchy for all within- and across-session parameters for all consonantal /r/ targets. Lilianne's accuracy was thus re-evaluated in accordance with the performance-based criterion. Based on the treating clinician's online ratings, Lilianne met the discharge criterion of >90% accuracy across targets at both word and sentence levels. Thus, no further treatment was provided, and this assessment was treated as the first post-treatment maintenance probe measure.

Word and sentence probes

Figures 2 and 3 depict Lilianne's accuracy in producing /r/ during baseline (BL), treatment (TX), and maintenance (MN)/follow-up intervals, with the treated interval shaded grey. In Figure 2, the

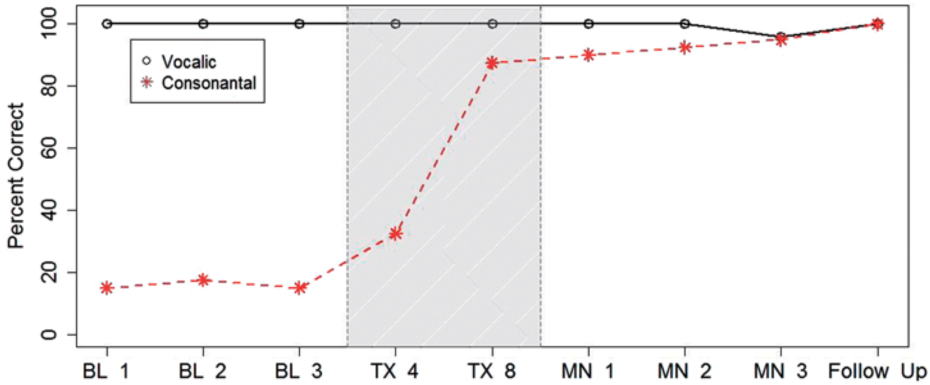

Figure 2. Lilianne's accuracy in producing untreated /r/ words elicited without biofeedback in the current study. *X*-axis represents probe elicitation session (BL = baseline, TX = within-treatment, MN = maintenance). *Y*-axis represents per cent of tokens rated perceptually correct based on mode across three blinded clinician listeners.

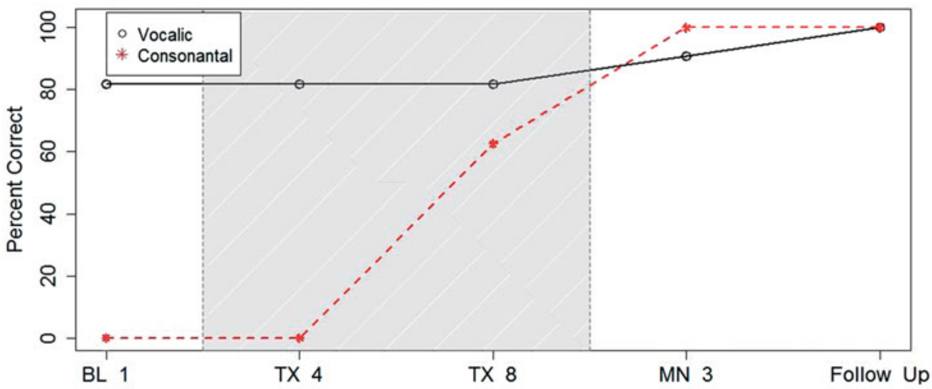

Figure 3. Lilianne's accuracy in producing /r/ in untreated sentences elicited without biofeedback. *X*-axis represents probe elicitation session (BL = baseline, TX = within-treatment, MN = maintenance). *Y*-axis represents per cent of /r/ words in sentence-level utterances rated perceptually correct based on mode across three blinded clinician listeners.

Y-axis represents the percentage of items in each /r/ word probe that were rated correct based on the mode across three blinded listeners. Figure 3 provides the same information for /r/ words elicited in the sentence probe measure. Vocalic /r/ variants are depicted with circles and a solid line, and consonantal variants are shown with asterisks and a dotted line.

For vocalic /r/ at the word level, Lilianne's productions were judged to be 100% accurate throughout the baseline phase. Her accuracy was much lower for consonantal /r/ at the word level, with a mean of 15.8% correct across baseline sessions. At the sentence level, she produced vocalic /r/ variants with 81% accuracy and consonantal variants with 0% accuracy at baseline.

Within treatment, Lilianne's accuracy in producing vocalic /r/ at the word level remained unchanged at 100% correct. Consonantal /r/ at the word level increased to 32.5% and then 87.5% correct. On sentence-level probes, Lilianne's accuracy for vocalic /r/ remained unchanged at 81% throughout the treatment interval, whereas consonantal /r/ increased to 62.5% correct.

In the three maintenance probes and the follow-up probe, Lilianne's accuracy for vocalic /r/ variants at the word level remained effectively at ceiling (mean 98.6% correct). Interestingly, her accuracy in producing consonantal /r/ at the word level increased steadily over the maintenance

Table 4. Standardised and unstandardised effect sizes for changes in accuracy at the word level from baseline to maintenance period.

/r/ Target	Standardised ES	Unstandardised ES
Vocalic – word level	−0.82	−1.4 percentage points
Consonantal – word level	37.6	76.7 percentage points
Vocalic – sentence level		9.1 percentage points
Consonantal – sentence level		100 percentage points

period and reached its highest point, 100% correct, in the 1-month follow-up session. At the sentence level, Lilianne's productions of vocalic /r/ were judged to be 91% accurate in maintenance session 3 and 100% accurate in the follow-up session. Consonantal /r/ at the sentence level was judged to be 100% accurate in both the maintenance and follow-up session.

Effect sizes comparing Lilianne's accuracy of /r/ production at the word level in the baseline versus the maintenance period are reported in Table 4. For vocalic variants, there was a small negative effect size, reflecting a ceiling effect. For consonantal /r/, a very large standardised effect size of 37.6 was recorded. The raw mean difference revealed a similarly striking change of 76.7 percentage points. At the sentence level, standardised effect sizes could not be calculated because the sentence probe was administered in only one baseline and one maintenance session. Raw mean differences revealed an increase of roughly 10 percentage points for vocalic /r/ and 100 percentage points for consonantal /r/ at the sentence level.

Discussion

Generalisation and challenge point principles

The participant in this single case study achieved strong and lasting gains across /r/ variants at both word and sentence levels. For vocalic /r/, gains were evident only at the sentence level due to ceiling effects at the word level. Consonantal /r/ showed improvement with very large effect sizes at both the word and sentence levels. Gains were fully maintained at the 1-month follow-up probe, and in fact, Lilianne's performance was suggestive of continued improvement in the absence of treatment.[2] Carryover to connected conversation was documented through elicitation of a 50-utterance spontaneous speech sample at the time of the 1-month follow-up session. The accuracy of /r/ sounds produced in this sample was evaluated by the same clinician who scored the pre-treatment conversational sample; individual utterances from both pre- and post-treatment samples were presented in random order without identifiers for blinded rating. The clinician's scoring indicated that at the 1-month follow-up session, Lilianne produced /r/ in connected conversation with 83% accuracy (34 out of 41 /r/ targets scored correct), an increase of 37 percentage points relative to her accuracy prior to treatment in the challenge point framework. Thus, these results provide preliminary support for the efficacy of a semi-structured implementation of the challenge point framework as a means of achieving generalisation and maintenance of treatment gains.

Through participant report and informal observation, we also found that Lilianne was far more motivated to participate in the treatment programme structured according to challenge point principles when compared to the previous study. Lilianne specifically reported that she

[2]Bacsfalvi and Bernhardt (2011) also reported observing continued improvement following the discontinuation of biofeedback intervention.

experienced a high level of satisfaction in connection with the varied, attainable intermediate goals presented within the challenge point treatment hierarchy. With opportunities to advance after every 10 trials, and with goals that were attainable from the participant's current level of functioning, Lilianne had numerous opportunities to experience the perception of skill mastery. This could enhance her intrinsic motivation in practice (Cameron, Pierce, Banko, & Gear, 2005), which could in turn help her maintain focus during the large number of trials required to master a new motor plan (e.g. Maas et al., 2008).

Limitations and future directions

The strength of the conclusions that can be drawn from the results reported here is intrinsically limited by the use of a single-subject case study design. It is not possible to generalise to the broader population based on data from a single individual whose response to treatment may have been unique in some unknown way. Moreover, the absence of a systematically manipulated control condition limits our ability to infer that the subject's gains over the study interval can be attributed to the treatment administered rather than some external factors.

Despite these threats to internal validity, there are several reasons to believe that Lilianne's /r/ production did improve as a consequence of the biofeedback treatment that she received across the previous study (McAllister Byun, Hitchcock, & Swartz, in press) and the present study. Although she received individual treatment targeting /r/ for 2.5 years prior to her enrollment in the first biofeedback study, Lilianne demonstrated a stable pattern of misarticulation affecting both vocalic and consonantal variants of /r/. She then made rapid gains over 8.5 weeks of biofeedback intervention, reaching virtually ceiling-level accuracy for vocalic /r/. Her pattern of mostly correct production of vocalic /r/ with mostly incorrect production of consonantal /r/ then remained stable until the onset of the current study. After biofeedback treatment resumed, Lilianne made rapid progress on consonantal /r/, as well as vocalic /r/ in the more complex sentence-level context. The fact that Lilianne's periods of progress coincided with the initiation and re-initiation of biofeedback treatment supports the hypothesis that her gains were caused by the treatment and not by maturation or other external factors. Finally, although there was no actively manipulated control condition, a naturally occurring control was provided by Lilianne's misarticulation of the untreated phonemes /s/ and /z/. In the maintenance and follow-up probes elicited at the end of the present study, Lilianne continued to misarticulate these sounds even at the single-word level. This provides additional informal evidence that her progress on /r/ is not attributable to maturation.

More limited evidence is provided by the present study on the clinically important question of whether biofeedback treatment structured according to challenge point principles is more effective in inducing generalisation than biofeedback on its own. We can begin to address this question by comparing Lilianne's progress on consonantal /r/ at the word level across the previous study, which provided 16 sessions of biofeedback treatment not structured according to challenge point principles, versus the present study, which provided 11 sessions of biofeedback treatment in the challenge point framework. (Only consonantal /r/ will be considered due to the ceiling effect limiting our ability to measure Lilianne's progress on vocalic /r/ in the current study). In the initial study, Lilianne's gains from baseline to maintenance yielded a standardised effect size of 4.0 (mean level difference of 34 percentage points), far smaller than the current effect size of 37.6 (mean level difference of 76.7 percentage points). While these results are suggestive, it is not possible to draw strong conclusions from these observations. Besides the fact that we report results from only a single participant, the design of this case study leaves open the possibility of an order effect or threshold effect. That is, it is possible that Lilianne's progress was slower during the initial period of biofeedback treatment because she needed to complete a certain duration of treatment or attain a threshold level of accuracy before entering into a more efficient phase of learning. There is thus a clear need to follow up on this research with controlled comparisons and a larger number of participants. Still, the difference in outcomes observed in this case study provides preliminary support for the effectiveness of challenge point methods in enhancing generalisation gains.

Conclusion

The findings from the present study support the growing body of research suggesting that visual biofeedback can facilitate correct production of /r/ in children whose errors do not respond to traditional methods. Further, the participant in this study successfully moved from the establishment phase of acquisition of the phoneme /r/ to generalisation and maintenance in the context of treatment incorporating a semi-structured implementation of the challenge point framework. Previous research in the framework of principles of motor skill learning has suggested that detailed qualitative feedback is effective for the initial establishment of a motor plan, but continued access to such input may inhibit generalisation to unlearned contexts (Newell et al., 1990). Drawing on challenge point principles, which suggest that optimal learning occurs when the level of difficulty is neither too high nor too low for the learner, this study incorporated progressive fading of biofeedback within a multilevel hierarchy that also manipulated target difficulty, level of clinician support, and order of target elicitation. The successful generalisation demonstrated by this case study participant encourages further investigation of the use of a semi-structured biofeedback treatment programme based on challenge point principles. Broader documentation of the efficacy of such a programme could serve to facilitate more widespread adoption of biofeedback for establishment, generalisation, and maintenance of correct production in children with persistent errors affecting sounds like /r/.

Acknowledgements

The authors gratefully acknowledge the contributions of the following individuals: Olivia Bell and Kristen Shubeck (student assistants); and Meghan Fougere, Risa Battino, Sarah Carmody, and Heather Campbell (clinical data raters). Finally, we thank ‘‘Lilianne’’ and her family for their cooperation throughout the study.

Declaration of interest

This project was supported by NIH R03DC012883 to Tara McAllister Byun. Aspects of this research were presented at the annual convention of the American Speech-Language Hearing Association in Chicago (2013).

References

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

Bacsfalvi, P. (2010). Attaining the lingual components of /r/ with ultrasound for three adolescents with cochlear implants. *Revue Canadienne d’Orthophonie et d’Audiologie*, *34*, 206–217.

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

Bacsfalvi, P., Bernhardt, B. M., & Gick, B. (2007). Electropalatography and ultrasound in vowel remediation for adolescents with hearing impairment. *Advances in Speech–Language Pathology*, *9*, 36–45.

Becker, M., & Levine, J. (2010). Experigen: An online experiment platform. Retrieved April 2013, from http://github.com/tlozoot/experigen.

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

Bernhardt, B. M., Gick, B., Bacsfalvi, P., & Ashdown, J. (2003). Speech habilitation of hard of hearing adolescents using electropalatography and ultrasound as evaluated by trained listeners. *Clinical Linguistics & Phonetics*, *17*, 199–216.

Busk, P. L., & Serlin, R. (1992). Meta-analysis for single case research. In T. R. Kratochwill & J. R. Levin (Eds.), *Single-case research design and analysis: New directions for psychology and education* (pp. 187–212). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cameron, J., Pierce, W. D., Banko, K. M., & Gear, A. (2005). Achievement-based rewards and intrinsic motivation: A test of cognitive mediators. *Journal of Educational Psychology*, *97*, 641–655.

Fawcett, S., Bacsfalvi, P., & Bernhardt, B. M. (2008). Ultrasound as visual feedback in speech therapy for /r/ with adults with Down Syndrome. *Down Syndrome Quarterly*, *10*, 4–12.

Fletcher, S. G., Dagenais, P. A., & Critz-Crosby, P. (1991). Teaching consonants to profoundly hearing-impaired speakers using palatometry. *Journal of Speech and Hearing Research*, *34*, 929–943.

Gibbon, F. E., & Paterson, L. (2006). A survey of speech and language therapists' views on electropalatography therapy outcomes in Scotland. *Child Language Teaching and Therapy*, *22*, 275–292.

Gick, B., Bacsfalvi, P., Bernhardt, B. M., Oh, S., Stolar, S., & Wilson, I. (2008). A motor differentiation model for liquid substitutions: English /r/ variants in normal and disordered acquisition. *Proceedings of Meetings on Acoustics*, *1*, 1–9.

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

Kaderavek, J. N., & Justice, L. M. (2010). Fidelity: An essential component of evidence-based practice in speech-language pathology. *American Journal of Speech-Language Pathology*, *19*, 369–379.

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

Martin, N. A., & Brownell, R. (2005). *Test of auditory processing skills*. Novato, CA: Academic Therapy Publications.

McAllister Byun, T., & 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., & Swartz, M. T. (in press). Retroflex versus bunched in treatment for /r/ misarticulation: Evidence from ultrasound biofeedback intervention. To appear in: *Journal of Speech, Language, and Hearing Research*, *2*, 2.

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

McGowan, R. S., Nittrouer, S., & Manning, C. J. (2004). Development of [r] in young, midwestern, American children. *Journal of the Acoustical Society of America*, *115*, 871–884.

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

Modha, G., Bernhardt, B. M., Church, R., & Bacsfalvi, P. (2008). Case study using ultrasound to treat /r/. *International Journal of Language & Communication Disorders*, *43*, 323–329.

National Institute on Deafness and Other Communication Disorders. (1994). *National strategic research plan*. Bethesda, MD: Department of Health and Human Services.

Newell, K. M., Carlton, M. J., & Antoniou, A. (1990). The interaction of criterion and feedback information in learning a drawing task. *Journal of Motor Behavior*, *22*, 536–552.

Preston, J., Brick, N., & Landi, N. (2013). Ultrasound biofeedback treatment for persisting childhood apraxia of speech. *American Journal of Speech-Language Pathology*, *22*, 627–643.

Ruscello, D. M. (1995). Visual feedback in treatment of residual phonological disorders. *Journal of Communication Disorders*, *28*, 279–302.

Rvachew, S., & Brosseau-Lapré, F. (2012). *Developmental phonological disorders: Foundations of clinical practice*. San Diego, CA: Plural Publishing.

Schmidlin, S., & Boyce, S. (2010, unpublished). /r/ pre and post test. *University of Cincinnati*.

Shawker, T. H., & Sonies, B. C. (1985). Ultrasound biofeedback for speech training. Instrumentation and preliminary results. *Investigative Radiology*, *20*, 90–93.

Shriberg, L. D. (1994). Five subtypes of developmental phonological disorders. *Clinics in Communication Disorders*, *4*, 38–53.

Shriberg, L. D., Austin, D., Lewis, B. A., McSweeny, J. L., & Wilson, D. L. (1997). The percentage of consonants correct (PCC) metric: Extensions and reliability data. *Journal of Speech, Language, and Hearing Research*, *40*, 708–722.

Shuster, L. I., Ruscello, D. M., & Smith, K. D. (1992). Evoking [r] using visual feedback. *American Journal of Speech-Language Pathology*, *1*, 29–34.

Shuster, L. I., Ruscello, D. M., & Toth, A. R. (1995). The use of visual feedback to elicit correct /r/. *American Journal of Speech-Language Pathology*, *4*, 37–44.

Smit, A. B., Hand, L., Freilinger, J. J., Bernthal, J. E., & Bird, A. (1990). The Iowa Articulation Norms Project and its Nebraska replication. *Journal of Speech and Hearing Disorders*, *55*, 779–798.

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

Vygotsky, L. S. (1978). *Mind and society: The development of higher mental processes*. Cambridge, MA: Harvard University Press.