

Tutorial

Tutorial: Using Visual–Acoustic Biofeedback for Speech Sound Training

Elaine R. Hitchcock,^a  Laura C. Ochs,^a  Michelle T. Swartz,^a  Megan C. Leece,^b 
Jonathan L. Preston,^b  and Tara McAllister^c 

^aMontclair State University, NJ ^bSyracuse University, NY ^cNew York University, NY

ARTICLE INFO

Article History:

Received April 29, 2022

Revision received August 26, 2022

Accepted October 1, 2022

Editor-in-Chief: Katherine C. Hustad

Editor: Ignatius Nip

https://doi.org/10.1044/2022_AJSLP-22-00142

ABSTRACT

Purpose: This tutorial summarizes current practices using visual–acoustic biofeedback (VAB) treatment to improve speech outcomes for individuals with speech sound difficulties. Clinical strategies will focus on residual distortions of /r/.

Method: Summary evidence related to the characteristics of VAB and the populations that may benefit from this treatment are reviewed. Guidelines are provided for clinicians on how to use VAB with clients to identify and modify their productions to match an acoustic representation. The clinical application of a linear predictive coding spectrum is emphasized.

Results: Successful use of VAB requires several key factors including clinician and client comprehension of the acoustic representation, appropriate acoustic target and template selection, as well as appropriate selection of articulatory strategies, practice schedules, and feedback models to scaffold acquisition of new speech sounds.

Conclusion: Integrating a VAB component in clinical practice offers additional intervention options for individuals with speech sound difficulties and often facilitates improved speech sound acquisition and generalization outcomes.

Supplemental Material: <https://doi.org/10.23641/asha.21817722>

A growing body of research has supported increased clinical use of visual biofeedback tools for remediation of speech sound deviations, particularly distortions affecting American English rhotics (Bacsfalvi et al., 2007; Bernhardt et al., 2005; Gibbon & Paterson, 2006; Hitchcock et al., 2017; McAllister Byun, 2017; McAllister Byun & Campbell, 2016; McAllister Byun et al., 2014, 2017; McAllister Byun & Hitchcock, 2012; Preston et al., 2013, 2014; Schmidt, 2007; Shuster et al., 1992, 1995; Sugden et al., 2019). Visual biofeedback offers a unique supplement to traditional treatment due to the inclusion of a visual representation of the speech sound, which can be used to make perceptually subtle aspects of speech

visible (Volin, 1998). As a result, the learner can alter their speech production by attempting to match a representation of an accurate target displayed in an external image. This tutorial summarizes the literature and describes clinical application of one type of biofeedback, visual–acoustic biofeedback (VAB).

Past research has shown that individuals with speech sound distortions who show a limited response to traditional interventions may benefit from therapy incorporating visual biofeedback (e.g., McAllister Byun & Hitchcock, 2012; Preston et al., 2019). If the speaker has a poorly defined auditory target, he/she may have difficulty imitating a clinician's auditory model but may succeed in matching a clearly defined visual representation of the target speech sound. Instead of relying on internal self-perception, clients are instructed to use the external image to gain insight into articulatory (i.e., ultrasound and electropalatography) or acoustic (spectrographic/spectral) information that is otherwise difficult to explain or teach. Furthermore, research exploring motor learning in nonspeech tasks has shown increased skill accuracy and

Correspondence to Elaine R. Hitchcock: hitchcocke@montclair.edu.

Disclosure: Elaine R. Hitchcock is a 2017 recipient of a PENTAX Medical Research Grant award. Tara McAllister oversees the direction of the *staRt* app for visual–acoustic biofeedback. Tara McAllister has an ownership stake in Sonority Labs LLC, a small business that has licensed the *staRt* software for commercialization. The other authors have declared that no other competing financial or nonfinancial interests existed at the time of publication.

less variability when an external focus of attention was incorporated into an oral movement task (Freedman et al., 2007). Thus, adopting an external focus of attention for speech movements may enhance retention of learned motor skills (Maas et al., 2008).

Previous literature has devoted considerable attention to types of biofeedback that provide a visual representation of the articulators (e.g., Gibbon et al., 1993; Preston et al., 2018; Sugden et al., 2019). Relatively less attention has been afforded to biofeedback of the visual–acoustic type, which may be more accessible due to the reliance on just an acoustic signal. This tutorial will briefly review the concepts of resonance and formants, define the characteristics of VAB, and describe different types of acoustic biofeedback. We will emphasize the use of a linear predictive coding (LPC) spectrum, as it is the most commonly used form of VAB in the research literature to date. We will briefly summarize several populations that may benefit from VAB. We will then describe its use for children with speech sound disorders, specifically rhotic errors, followed by an example of how to implement treatment (McAllister Byun, 2017; McAllister Byun & Hitchcock, 2012; McAllister Byun et al., 2014, 2016). Lastly, as an example of clinical utility, we will provide a detailed description of rhotic acquisition training using VAB while integrating specific cuing strategies described in previous literature focused on traditional approaches to /r/ treatment (Preston et al., 2020). This tutorial will demonstrate how clients can be taught to identify and modify their productions to match an acoustic representation or *formant pattern* that characterizes the speech target. We offer guidelines for intervention using VAB with the long-term goal of improving outcomes for challenging speech sound distortions.

What Is VAB?

Unlike other forms of visual biofeedback that reveal articulatory actions, VAB depicts an *acoustic representation* of the speech sound, such as an LPC spectrum or spectrogram, which will be discussed in further detail in the following sections (see Figure 1 for an example using an LPC spectrum). The client and clinician together view a real-time dynamic image of the speech signal simultaneously paired with a template representing the acoustic characteristics of the target sound. The client can be cued to adjust their speech output to match a pre-selected template or target on the screen (see Figure 2), which can be paired with cues for articulator placement. By watching the visual display change in response to their articulatory changes, learners can build stronger associations between articulatory postures and auditory-acoustic outcomes (Awan, 2013). At the same time, because different articulatory configurations can yield similar auditory and acoustic consequences, VAB allows the learner to find their own articulatory approach to a given target sound. This level of flexibility may be beneficial in the context of sounds like /r/ (McAllister Byun et al., 2014), which can be realized with a range of articulatory configurations, as detailed below.

Resonance and Formants

Briefly summarized, *resonance* occurs when the molecules of air in the vocal tract vibrate in response to the sound source, namely, vibratory behavior of the vocal folds. Depending on the size and shape of the supraglottic cavities (pharyngeal and oral cavities), specific frequencies will be amplified or attenuated (Ladefoged, 1996). *Formants* are those frequencies that are amplified because

Figure 1. Sample linear predictive coding (LPC) spectrum showing oral and pharyngeal cavity during production of /i/ (“ee”). We have labeled F3 for clarity; however, the locations of F1 and F2 are typically sufficient for identifying vowels.

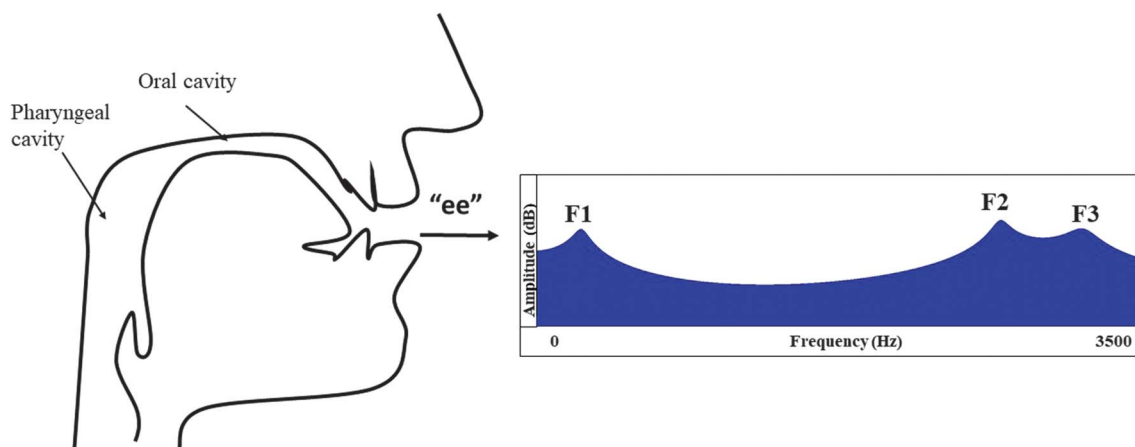
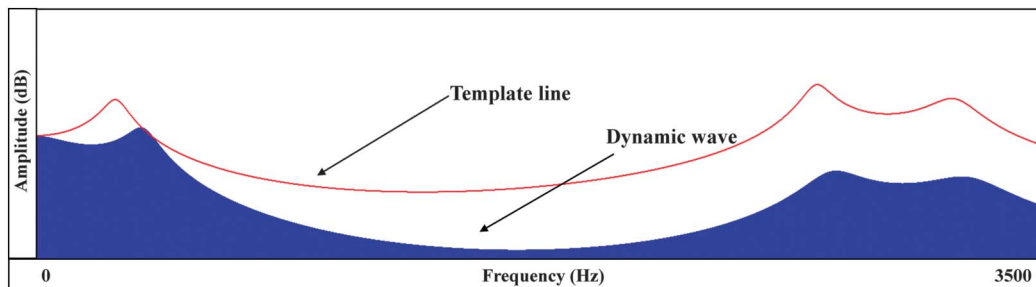


Figure 2. Sample linear predictive coding (LPC) spectrum of /i/ with target template.



they align with the natural resonant frequencies of the vocal tract (Fant, 1960). In general, a large cavity volume will resonate at a low frequency, whereas a small cavity volume will resonate at a high frequency (Zemlin, 1998). In a simplified two-container model of vocal tract resonance, the size and shape of the pharyngeal and oral cavities determine the first (F1) and second (F2) formant frequencies, respectively. The frequency of F1 is determined primarily by tongue body height (vertical plane), which affects the volume of the pharyngeal cavity. Specifically, the pharyngeal cavity will contain a large volume of air when the tongue is high in the mouth, pulling the tongue root up and out of the pharyngeal space; its volume is smaller when the tongue body is lower. This results in low F1 for a high vowel (e.g., /i/) and higher F1 values for a low vowel (e.g., /a/; Peterson & Barney, 1952). The frequency of F2 is determined by tongue position in the anterior–posterior plane, which affects the volume of the oral cavity in front of the point of maximal tongue constriction. When the tongue is anterior in the mouth (as in /i/ or /e/), this space is small, resulting in a high F2; when the tongue is posterior (as in /u/ or /o/), this space is larger, resulting in a low F2. Thus, a high front vowel such as /i/ is characterized by a low F1 and a high F2 (see Figure 1). When a speaker is attempting to produce an /i/ sound, more complex acoustic dynamics are at play. F1 and F2 are typically located in a neutral or central position of the vowel space, similar to their position for schwa. The acoustic hallmark of an accurate /i/ production is a low third formant (F3), which is associated with supraglottic vocal tract constrictions at the lips, anterior oral cavity, and pharyngeal region (Chiba & Kajiyama, 1941). Constrictions in these regions can be formed with a range of tongue configurations, as we describe in more detail below (Delattre & Freeman, 1968; Tiede et al., 2004; Zhou et al., 2008). In this context, speakers must learn what configuration of their own articulatory structures best maps to the auditory-acoustic signature of /i/ (e.g., Guenther et al., 1998). Displaying acoustic information may be particularly useful in order to allow exploration of articulatory-acoustic mappings in a therapeutic context.

LPC Spectrum Versus Spectrogram

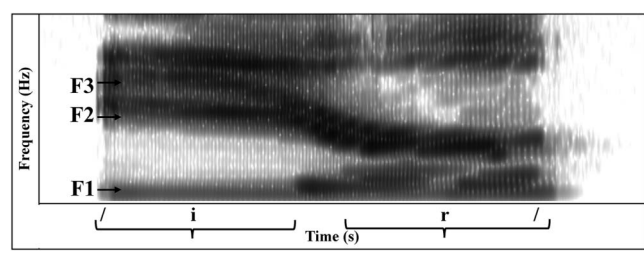
Altering the shape of the vocal tract changes its resonant characteristics, which in turn changes the frequency of the formants. We can visualize these frequency changes in various forms, most notably with an LPC spectrum or spectrogram (see Figure 1). When viewing an LPC spectrum, frequency, typically measured in Hertz (Hz), is represented on the *x*-axis whereas amplitude, typically measured in decibels (dB), is represented on the *y*-axis. An LPC spectrum allows visualization of the amplitude of the frequency components of a speech sound; formants or resonant frequencies of the vocal tract are visualized as vertical peaks in the frequency range of the spectrum. Traditional LPC spectra do not include a time axis, thus reflecting a static image of selected point in time within a speech sound. However, current technologies allow for display of dynamic, real-time components of the changing LPC spectrum. In several existing software programs for acoustic analysis, speakers may view changing formant values as moving peaks corresponding to articulatory modification of the vocal tract (see Supplemental Material S1).

In comparison, a spectrogram allows the learner to visualize the relative amplitudes of all frequency components of a sound, with time represented on the *x*-axis and frequency on the *y*-axis. The amplitude of different frequency components is represented with gradations of color or darkness. The formant frequencies, which are characterized by high amplitude, appear as dark horizontal bands that shift up and down with the changing resonances of the moving vocal tract (see Figure 3). This display of acoustic information can be used in a clinical context to teach clients to modify their speech sounds by seeing changes in format patterns (Shuster et al., 1995).

Who Can Benefit From VAB?

Several populations, including second language learners, individuals who are hard of hearing, and individuals who have been diagnosed with speech sound disorder, are uniquely suited to benefit from spectrographic or spectral VAB due to the nature of the target adopted in therapy or

Figure 3. Spectrogram showing an adult speaker production of “ear.” The shifting formants are reflected in the changing horizontal bands.



training (Akahane-Yamada et al., 1998; Brady et al., 2016; Carey, 2004; Crawford, 2007; Dowd et al., 1998; Ertmer & Maki, 2000; Ertmer et al., 1996; Kartushina et al., 2015, 2016; Li et al., 2019; Maki & Streff, 1978; Olson, 2014; Stark, 1971). These populations demonstrate speech characteristics that warrant an initial focus on the segmental aspects of speech (vowels/consonants). A visual-acoustic intervention program can be effective when targeting sonorant speech sounds, where the speaker’s attempts can be lengthened and compared with a preselected target or template. While fricatives and suprasegmental factors (intonation, stress, and duration) may also be identified as potential targets for which different types of VAB are appropriate (Swartz et al., 2018; Utianski et al., 2020), such approaches are beyond the scope of this tutorial. Last, our anecdotal observations suggest that individuals age 8 years or older tend to benefit most from VAB, presumably because VAB requires that the client can comprehend a dynamic visual display and connect that display with their articulatory changes, a task that can be challenging for younger children. However, it is possible that some younger children may benefit this technique. Careful consideration of factors such as attention and motivation may be more important than age when determining candidacy for potential VAB clients.

Individuals With Hearing Loss

Several studies have suggested that spectrographic biofeedback can be successfully utilized as a speech training tool for individuals who are deaf or hard of hearing (see Table 1; Crawford, 2007; Ertmer & Maki, 2000; Ertmer et al., 1996; Maki & Streff, 1978; Stark, 1971). It is well known that the severity of hearing loss impacts intelligibility of speech. Furthermore, children with severe–profound hearing loss often demonstrate poorer speech intelligibility compared with children with mild–moderate hearing loss. Speakers who are deaf or hard of hearing may experience challenges with segmental aspects of speech production (consonants, vowels, and diphthongs) and suprasegmental aspects of speech (Culbertson & Kricos, 2002).

In children with hearing loss, the use of VAB can compensate for reduced auditory access to the full range

of the speech spectrum. VAB offsets the lack of salient auditory cues by providing visual representations of the speech signals. VAB circumvents an impaired auditory feedback mechanism, instead allowing the client to “see” what happens to the target sound in response to articulatory changes during training of vowels (Ertmer et al., 1996; Maki & Streff, 1978; Stark, 1971) and/or consonants (Ertmer & Maki, 2000; Shuster et al., 1992).

Second Language Learners

Research incorporating various forms of VAB (LPC spectra, spectrograms, and vowel charts) has been successfully implemented in speech training programs with adult learners who aim to master speech sounds in a second language or L2 (Akahane-Yamada et al., 1998; Brady et al., 2016; Carey, 2004; Dowd et al., 1998; Kartushina et al., 2015, 2016; Li et al., 2019; Olson, 2014). A number of studies have shown progress in the acquisition of L2 speech sounds after a relatively short period of VAB training (see Table 1).

Children With Speech Sound Disorders

A growing body of research suggests that technology-enhanced interventions such as VAB could improve outcomes for children with challenging speech sound distortions, notably distortions of /ɹ/ (see Table 2 for a detailed list of studies); we will draw on the example of rhotic biofeedback throughout this tutorial as a demonstration of clinical application of VAB. American English rhotics are among the most frequently misarticulated sounds and are widely acknowledged as some of the hardest to treat, often proving resistant to traditional therapy techniques (Ruscello, 1995; Shuster et al., 1995).¹ In the following sections, we will summarize the articulatory complexity of the /ɹ/ sound and explain the unique benefits of VAB for this population.

Articulation of American English rhotics. Both the high prevalence and treatment-resistant nature of American English rhotic errors are commonly attributed to the articulatory complexity of an accurate production (Gick et al., 2007). Typically, /ɹ/ is produced with two major lingual constrictions, one anterior and one posterior in the vocal tract (e.g., Delattre & Freeman, 1968). The posterior constriction is characterized by tongue root retraction to narrow the pharyngeal cavity (Boyce, 2015). Numerous investigations have shown that the anterior tongue configuration for /ɹ/ is subject to variability, both across and within speakers (Delattre & Freeman, 1968; Tiede et al., 2004; Zhou et al., 2008). Two major variants are commonly

¹Distortions of /s/ represent another common clinical challenge and are also amenable to remediation with VAB. However, here, we focus on /ɹ/, because it is much better represented than /s/ in the literature to date.

Table 1. Selected studies using visual–acoustic biofeedback (VAB) for speech training outside the context of speech sound disorder.

Study	Year	N	Sex	Age	Population	Target	Type of biofeedback	Duration	Total sessions
Crawford	2007	3	M-2; F-1	7–12	HoH	Consonants	Spectrogram	30 min, 1–2× per week	2–24
Ertmer et al.	1996	2	M-1; F-1	9	HoH	Vowels	Spectrogram	30 min, 3× per week	~60
Ertmer & Maki	2000	4	F-4	13;3–15;9 M = 14;9	HoH	/m/ /t/	Spectrogram	20 min 4× per week	8
Stark	1971	1	M-1	8;0	HoH	Consonants/ vowels	Spectrogram	Unknown	Unknown
Akahane-Yamada et al.	1998	10	M-9; F-1	18;0–24;0 M = 21	L2	/ɪ/ /I/	Spectrogram	~100 min 5 hr per tx	3
Brady et al.	2016	1	M-1	24;0	L2	Vowels	Spectrogram	25–30 min 2/3 per week	11
Kartushina et al.	2016	20	M-2; F-18	M = 21;9	L2	Vowels	Real-time F1/F2 chart	600 trials per target ^a	3
Kartushina et al.	2015	27	M-7; F-20	M = 24;8	L2	Vowels	Real-time F1/F2 chart	45 min 2/3 per week	5
Li et al.	2019	60 ^b	F-60	18;0–30;0	L2	Vowels	LPC	30-min	2

Note. M = male; F = female; HoH = hard of hearing; L2 = second language clients; LPC = linear predictive coding.

^aDuration dependent on completion of trials. ^bVisual–acoustic biofeedback randomly assigned to approximately half of sample.

Table 2. Summary of visual–acoustic biofeedback (VAB) studies in the context of speech sound disorder.

Study	Year	<i>N</i>	Sex	Ages	Biofeedback type	Dose frequency	Total sessions	No. of teaching episodes	CII	Effect size ^a (range)
McAllister Byun et al.	2016	9	M-6; F-3	6;8–13;3 (<i>M</i> = 10;0)	LPC	30 min 2× per week	~16	60	968	<i>d</i> ₂ –1.2 to 5.5
Benway et al.	2021	7	M-3; F-4	9;5–15;8 (<i>M</i> = 12;3)	LPC/ultrasound	101 min/2× per week	10	400	4000	<i>d</i> ₂ 0–3.11
McAllister Byun	2017	7	M-5; F-2	9;0–15;0 (<i>M</i> = 12;3)	Traditional/LPC	30 min 2× per week	20	60	1200	<i>d</i> ₂ –5.3 to 9.87
McAllister Byun & Hitchcock	2012	11	M-10; F-1	6;0–11;9 (<i>M</i> = 9;0)	Traditional/LPC	30 min 2× per week	20	60	1200	<i>d</i> (<i>group</i>) 0.32
Shuster et al.	1995	2	M-1; F-1	10;0–14;0 (<i>M</i> = 12;0)	Traditional/ spectrogram	(a) 50 min 2× per week (b) 60 min 1× per week	24	150 175	(a) 3600 (b) 1400	NR
Hitchcock et al.	2017	4	M-3; F-1	8;8–13;0 (<i>M</i> = 9;10)	LPC	60 min 2× per week	20	192	3840	<i>d</i> ₂ –1.63 to 18.92
McAllister Byun & Campbell	2016	11	M-7; F-4	9;3–15;10 (<i>M</i> = 11;3)	Traditional/LPC	2× per week ^b	20	60	1200	<i>d</i> ₂ –0.27 to 20.45
McAllister Byun et al.	2017	1	F-1	13;0	Traditional/LPC (staRt app)	30 min 1× per week	20	60	1200	NR
Peterson et al.	2022	4	M-2; F-2	9;0–10;3 (<i>M</i> = 9;8)	LPC (staRt app)/ telepractice	2–3× per week ^b	16	200	3200	<i>d</i> ₂ 5.3–67.7

Note. CII = cumulative intervention intensity; M = male; F = female; LPC = linear predictive coding; NR = not reported.

^aEffect sizes were calculated for each individual by comparing /ɹ/ productions elicited in single-word probes administered during the true baseline phase (prior to the initiation of any treatment) and the posttreatment maintenance phase. Effect sizes were standardized using Busk and Serlin's (1992) *d*² statistic (Beeson & Robey, 2006), which pools *SD*s across baseline and maintenance periods to reduce the number of cases where effect size cannot be calculated due to zero variance at baseline. ^bDuration of session not reported.

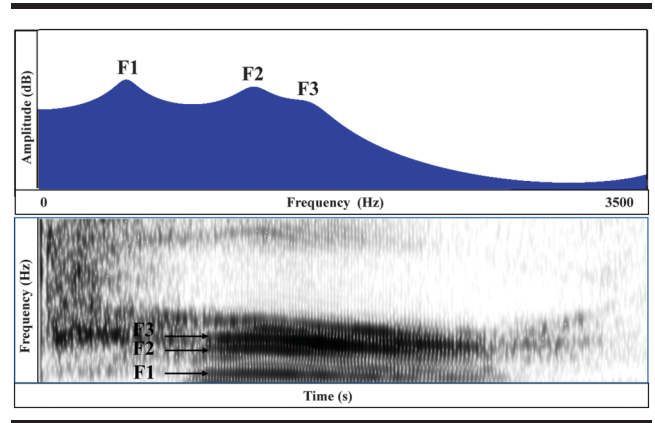
reported: the retroflex variant of /ɹ/, where the tongue tip is raised and may be curled up slightly at a point near the alveolar ridge, and the bunched variant of /ɹ/, where the tongue tip is lowered and the anterior tongue body is raised to approximate the palate. However, there is a great deal of variability between these extremes, and some tongue shapes do not fit well into either category (Boyce, 2015). An added complication is that many speakers use different tongue shapes across different phonetic contexts (Mielke et al., 2016; Stavness et al., 2012). Many speakers also produce /ɹ/ with slight labial constriction (King & Ferragne, 2020). Importantly, the various articulatory configurations for /ɹ/ appear to result in relatively consistent acoustic patterns at the level of the first three formants and are, to the best of our knowledge, perceptually equivalent (Zhou et al., 2008).

In summary, treatment for misarticulation of /ɹ/ involves cueing the learner to imitate tongue constrictions that are complex and vary across speakers and contexts, making it hard for the clinician to know which rhotic variant to cue. It poses a further challenge because the crucial tongue constrictions are contained within the oral cavity and, as such, cannot be visualized without some form of instrumentation. Finally, because /ɹ/ is produced with limited contact between articulators, there is little tactile feedback to support learners in achieving the desired tongue configurations. Although, in some cases, traditional motor-based treatment strategies can successfully help remediate rhotic errors (see Preston et al., 2020, for a full discussion), it is common for clients to demonstrate limited progress, which may result in frustration or even termination of treatment (Ruscello, 1995).

VAB for American English rhotics. Regardless of the articulatory complexity of the American English /ɹ/, the sound that is produced yields a distinctive formant pattern that makes it particularly suitable for treatment with VAB. A lowered third formant (F3), occasionally low enough that it appears to merge with the second formant (F2), is considered the hallmark of American English /ɹ/ (e.g., Boyce & Espy-Wilson, 1997). In contrast, distortions of /ɹ/ are characterized by F3 values between 2500 and 3500 Hz, compared with correct /ɹ/ productions typically lower than 2500 Hz for children and 2000 Hz for adults (Campbell et al., 2017; Lee et al., 1999; Shriberg et al., 2001). Figure 4 shows the close spacing of F2 and F3 in an LPC spectrum of syllabic /ɹ/. With VAB, clinicians can use the stable acoustic signature of /ɹ/ to help the client identify the movements of the tongue that result in acoustic changes in the direction of a more accurate /ɹ/ sound.

As noted above, because /ɹ/ can be produced with many different articulatory postures (e.g., bunched and retroflex), it can be challenging to identify optimal articulatory positions for some clients. Thus, the fact that VAB emphasizes the acoustic target rather than a specific tongue

Figure 4. Linear predictive coding (LPC) spectrum and spectrogram of adult correct /ɹ/ generated in the Sona-Match module of the Computerized Speech Lab (PENTAX Medical, Model 4500).



configuration represents a particular strength in the context of treatment for /ɹ/ misarticulation. Instead, the clinician is free to select any cues and feedback that they judge will help the client get closer to the desired acoustic characteristics of /ɹ/ (McAllister Byun et al., 2016). As we describe below, however, we recommend pairing visual feedback with articulatory cues focused on achieving an adequate oral constriction, lowering the tongue dorsum/body, elevating the lateral margins of the tongue, retracting the tongue root into the pharyngeal cavity, and achieving slight lip constriction (see Preston et al., 2020, for details). The opportunity to observe incremental acoustic changes in connection with their articulatory adjustments can help clients acquire motor routines for postures that offer little tactile or kinesthetic feedback (McAllister Byun & Hitchcock, 2012).

Not only does the articulatory complexity of /ɹ/ make it a difficult sound to remediate when produced in error, but many children presenting with /ɹ/ misarticulations also lack the auditory acuity to recognize rhotic errors in their own speech. Shuster (1998) reported that children with /ɹ/ misarticulation showed a decreased ability to discriminate correct versus distorted /ɹ/ sounds in their own output, making it harder to benefit from treatment in which the clinician supplies an auditory model of the target sound and prompts the child to match it. Similar results were reported by Cialdella et al. (2021) and Hitchcock et al. (2020), whose findings showed that typically developing children demonstrated more consistent classification of items along a synthetic continuum from /ɹ/ - /w/ compared with children with rhotic speech errors.

Evidence base for VAB in treatment of rhotic misarticulation. Both spectrograms and LPC spectra depict the formants or resonant frequencies of the vocal tract, which appear as horizontal bars in the former and vertical peaks in the latter. The early visual-acoustic literature focused on the use of spectrograms to teach clients to recognize and

attempt to match the formant pattern characterizing a target sound (Shuster et al., 1992; Shuster et al., 1995). More recent research has focused on the use of real-time LPC spectra generated either with Sona-Match or staRt software, described below. While both a spectrogram and LPC spectrum display formant information, recent studies have favored the use of the LPC spectrum as the acoustic biofeedback modality, because the display is visually less complicated than a spectrogram and, as a result, potentially less challenging for children to interpret changing formant patterns (see Figure 4).

Early case studies using spectrograms provided meaningful foundational evidence for the use of VAB in the treatment of /ɹ/ for children with speech sound disorder. Shuster et al. (1992, 1995) found that spectrograms can be used by clinicians to support effective intervention for residual rhotic misarticulation. They reported successful implementation of a spectrographic biofeedback program for three individuals with speech sound disorder, ages 10, 14, and 18 years. In these two small-scale studies, participants were described as nonresponders after receiving at least 2 years of traditional articulation treatment. Before the start of VAB intervention, all participants demonstrated 0% accuracy in /ɹ/ productions. After two to six sessions of spectrographic biofeedback intervention, all participants had attained at least 70% correct productions of isolated sustained /ɹ/. By the 11th session, all participants were producing /ɹ/ in isolation and rhotic diphthongs with 80%–100% accuracy. Generalization to spontaneous conversation and sentence-level utterances was reported for the 10- and 14-year-old participants, respectively (Shuster et al., 1995).

Recently, several small-scale experimental studies have indicated that VAB real-time LPC spectra can also represent an effective form of intervention for residual /ɹ/ distortions (McAllister Byun, 2017; McAllister Byun & Campbell, 2016; McAllister Byun & Hitchcock, 2012; McAllister Byun et al., 2016). McAllister Byun and Hitchcock (2012) investigated the efficacy of VAB using a single-case experimental design in which participants were transitioned from traditional motor-based treatment to spectral biofeedback in a staggered fashion after 4–6 weeks. They found that eight of 11 participants (ages 6;0–11;9 [years;months]) showed clinically significant improvement over the 10-week course of treatment and, in six of these eight participants, gains were observed only after the transition to VAB treatment. In another single-case experimental study, McAllister Byun et al. (2016) found that six of nine participants with residual rhotic errors demonstrated sustained improvement on at least one treated rhotic target after 8 weeks of VAB, despite previously showing no success over months to years of traditional articulatory intervention. A subsequent single-case experimental study of 11 children who received both traditional and biofeedback

treatment in a counterbalanced order (McAllister Byun & Campbell, 2016) revealed a significant interaction between treatment condition and order, such that individuals who received a period of VAB followed by a period of traditional treatment showed significantly greater accuracy on generalization probes than individuals who received the same treatments in the reverse order. McAllister Byun (2017) conducted a single-case randomization study in which seven participants received both spectral biofeedback and traditional treatment in an alternating fashion over 10 weeks of treatment. In that study, three participants showed significantly greater within-session gains in biofeedback than traditional sessions, with no participant showing a significant difference in the opposite direction (McAllister Byun, 2017). Peterson et al. (2022) used a similar single-case randomization design in a study of treatment using staRt biofeedback software (described below) via telepractice. In that study, all four participants made substantial gains in rhotic production accuracy, and one participant showed a significant advantage for biofeedback over traditional treatment (although interpretation of this result was complicated by the fact that random assignment resulted in a concentration of biofeedback sessions in the early stages of treatment). Finally, Benway et al. (2021) treated 9- to 16-year-old children with /ɹ/ errors using both VAB and ultrasound biofeedback in an alternating fashion and found that five of seven children showed evidence of acquiring /ɹ/. Only one participant showed a significant difference in within-session acquisition of /ɹ/ between the two treatment types, and this participant showed an advantage for VAB compared with ultrasound sessions. Overall, findings from these studies indicate that VAB intervention programs using an LPC spectrum can have positive outcomes for children with residual distortions of /ɹ/ who have not responded to traditional intervention approaches.

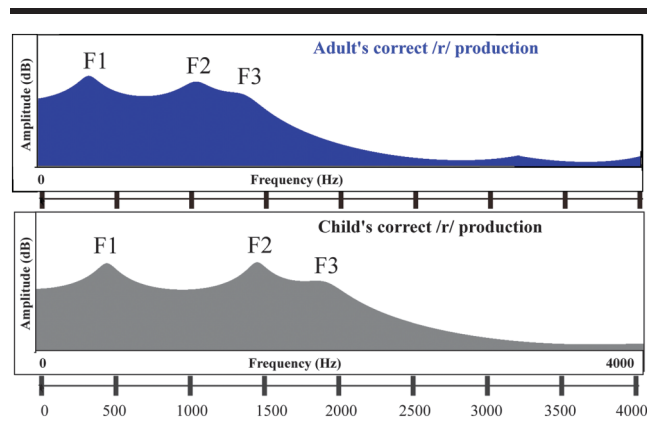
The remainder of this tutorial will provide suggested guidelines for the use of VAB to encourage acquisition of North American English /ɹ/. We will provide details on how to set an appropriate target, how to orient learners to the VAB display, how to integrate specific articulatory cueing strategies developed in the context of traditional treatment (Preston et al., 2020), and how to deploy these cues in accordance with the principles of speech-motor learning (Maas et al., 2008) to encourage generalization across speech contexts.

Suggestions for Implementing VAB Using the LPC Spectrum for /ɹ/ Training

Selecting an Appropriate LPC Target Pattern

VAB requires a target or template representing the desired acoustic output that the client can attempt to match during their practice productions. However, the

Figure 5. Adult and child correct linear predictive coding (LPC) spectrum or /ɹ/.



same target cannot be used for all clients because formant frequencies are influenced by vocal tract size.² That is, the placement of the individual formant peaks represented on the LPC spectrum differs depending on the size of the vocal tract, although the ratio of the distance between the peaks remains roughly the same (see Figure 5). Thus, the best target for a client would be selected from a typical speaker who is approximately the same age, size, and sex. Once the client begins to establish an approximation of a perceptually acceptable rhotic production, the clinician may choose to adjust the target (e.g., replacing a target derived from a different speaker with a target generated from the client's best approximation). A client's formant pattern targets may be refined over time as the client's rhotic accuracy improves, providing visual evidence of a progression toward age-appropriate production. Alternatively, the original target from a different speaker may be retained throughout the duration of treatment.

In the Computerized Speech Lab (CSL), Sona-Match module (PENTAX Medical, 2019, Model 4500), the LPC spectrum has three different viewing/window settings: child, adult female, and adult male. Selection is generally based upon the age and gender of the individual receiving the intervention (see Supplemental Material S2). The target used in VAB is a template or trace representing a snapshot of the spectral envelope of the entire formant signature for /ɹ/ (see Figure 2 and Supplemental Material S2).

²Selecting a target template is dependent upon the VAB software program. Regardless of the software selections, we suggest that clinicians first determine the approximate location of the third formant. Note that the location of F3 in a misarticulated /ɹ/ is expected to be somewhere around 3500 in a younger child (9 years old and under) and 3000 in an older child (10 years and up). In a correct /ɹ/, F3 is expected to be somewhere around 2000 Hz (see Lee et al., 1999, for detailed breakdown by age level).

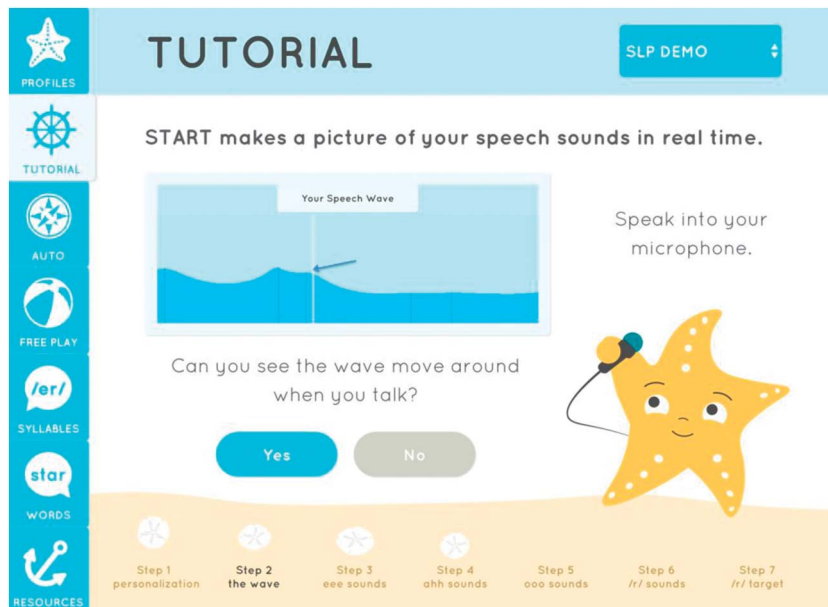
That is, although the focus in VAB for /ɹ/ is on the location (peak) of F3 and the distance between F2 and F3, a peak representing F1 is also visible in the target. Templates are not provided by default by the manufacturer; however, templates created by our team are available at Open Science Framework (<https://osf.io/kj4z2/>). By contrast, in the staRt app (McAllister Byun et al., 2017; BITS Lab NYU, 2020), the target takes the form of a single adjustable slider that is intended to represent the center frequency of F3, with the goal of drawing the client's attention to this critical peak in the display. When a user creates a profile on the staRt app, a target frequency is automatically generated based on the age and gender entered. The target is drawn from normative data representing typically developing children's productions of /ɹ/ (Lee et al., 1999). Users of the staRt app are instructed that they can adjust the slider as desired, but that target values below 1500 Hz are not recommended (see Figure 6).

Introductory Sessions (One to Two Sessions)

Establishing Comprehension of Acoustic-Articulatory Connection

Prior to initiating treatment, a key factor for success is the inclusion of one to two introductory sessions to ensure that the client understands how lingual movements modify the vocal tract and how those changes can be visualized via alterations of the formant peaks in the LPC spectrum. The potential for improvement may be compromised if the acoustic information in the LPC spectrum is not interpretable to the learner. To offset this possibility, an introductory session prior to the initiation of detailed practice is beneficial to provide an explanation of the bio-feedback spectrum. In this session, the clinician should demonstrate how lingual movements are reflected in the changing formant pattern using maximally opposing vowels, such as /i/ and /a/ to optimize the differences in the ratio of formant peaks (see Figure 7). Once the client has observed the clinical demonstration, they are encouraged to move their tongue around in their mouth while observing how their lingual movements shift the formants ("peaks" or "bumps") when different sounds are produced. A verbal comprehension check is recommended prior to moving forward to the treatment phase to ensure that the client adequately understands the relationship between articulatory changes and formant patterns. A version of the script used in our current clinical research using the CSL Sona-Match is included in Appendix A. (Interested clinicians may access a video example of an introduction to VAB located in our Supplemental Material S3). The staRt app has an interactive tutorial intended to guide the client and clinician through this basic information (as well as information about the acoustics of /ɹ/, described below); a video narration of the staRt tutorial is included in Supplemental Material S4.

Figure 6. Image of staRt app (BITS Lab NYU, 2020). Used with permission.



Characteristics of Spectral Changes During Articulatory Exploration

Once basic comprehension of the relationship between articulator changes and acoustic outputs is established, clients can then be familiarized with the concept of matching formant templates in a task involving sounds that they can articulate accurately. One strategy is to present the client with an unidentified vowel template and then direct them to guess which vowel the template represents.³ The client may be asked to produce several different vowels suggested by the clinician and then look for the LPC spectrum that is the closest match to the vowel template. Following several successful matches, the client is typically ready to progress to the next phase of treatment targeting rhotic productions.

Maximal benefit from spectral biofeedback stems from the client's ability to interpret changing formant patterns in real time. Initially, it is recommended that the clinician provide an age-appropriate verbal explanation of the formants, specifically highlighting that F2 and F3 (the second and third "bumps" or "peaks") are far apart in an incorrect /r/ sound but move close together or merge in correct /r/ production (see Figure 8 for an example of an LPC spectrum for incorrect and correct /r/). A comprehension check where the client is asked to

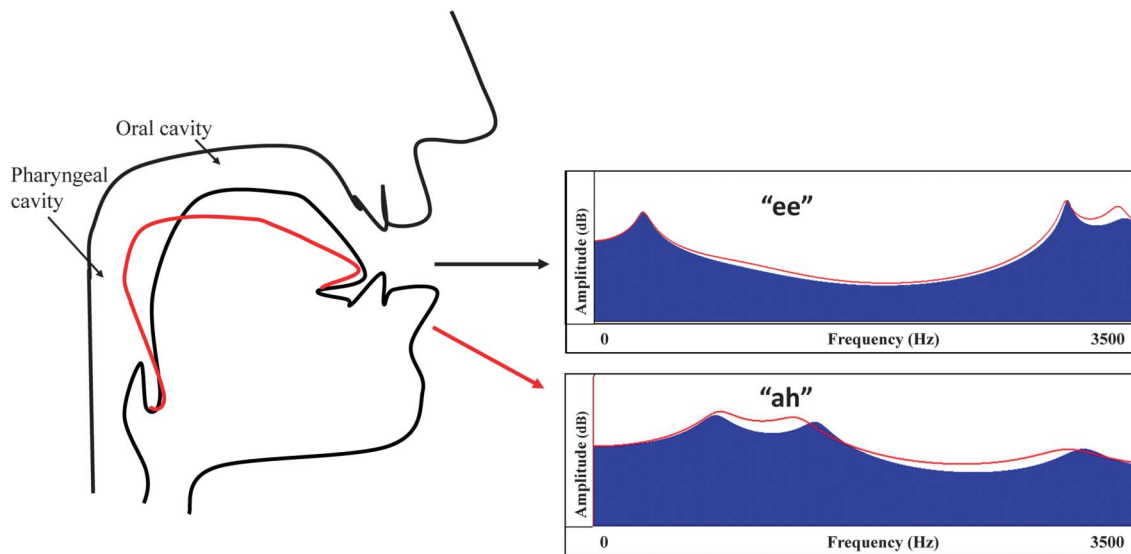
verbally describe the visual properties of correct and incorrect /r/ as reflected in the LPC spectrum and/or select an image depicting the requested target is helpful prior to initiating biofeedback treatment. Either static or dynamic images may be used; see Supplemental Material S5 for a teaching example of a transition from incorrect to correct /r/ production. Once they have been familiarized with the acoustic characteristics of correct and incorrect /r/, participants can be presented with an appropriate rhotic template superimposed over the dynamic LPC spectrum and encouraged to match the pattern of the formant peaks by modifying their vocal tract configurations. During treatment, the lowering of F3 may initially be a slight or gradual change, but learners can begin to associate this with a successive approximation of the target, enabling them to recognize when their productions are getting closer.

Articulatory Cues for Rhotic Sounds

While learners using biofeedback can be given general encouragement to explore a wide range of vocal tract shapes in an effort to achieve a closer match with the formant template, it is often judged clinically useful to pair the spectral biofeedback image with specific cues for articulator placement. Because part of the rationale for adopting biofeedback comes from its ability to engender an external direction of attentional focus, which has been associated with improved learning outcomes in nonspeech motor tasks (Maas et al., 2008), there is a possible theoretical argument whereby biofeedback could be rendered less

³As indicated previously, vowel templates are dependent on the VAB software program. Clinicians may also create and save additional vowel templates (directions provided in the SonaMatch User Manual) keeping in mind that the target template should be generated from a speaker who is relatively well matched for vocal tract size.

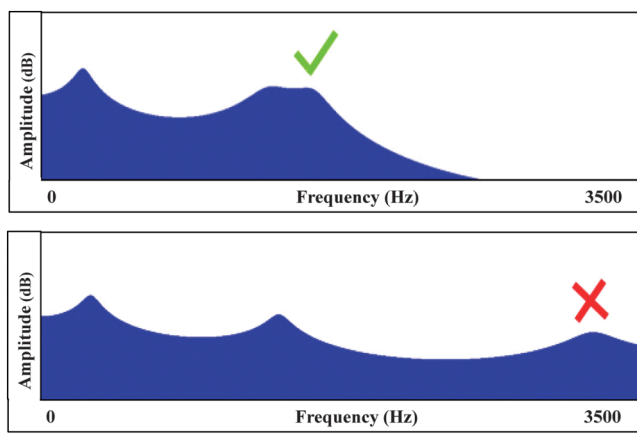
Figure 7. Linear predictive coding (LPC) images comparing /i/ (“ee”) and /a/ (“ah”).



effective through the incorporation of articulator placement cues that direct the learner’s attention inward. This question was investigated in the work of McAllister Byun et al. (2016), in which all participants received VAB treatment, half with explicit cues for articulator placement and half whose cues only referenced the visual–acoustic display. There were no significant differences in outcomes between the two attentional focus conditions, and the authors concluded that providers of VAB could incorporate any cues they judge clinically useful, including articulator placement cues. Recall that American English rhotics are typically characterized by three vocal tract constrictions: in the lips, anterior oral cavity, and posterior pharyngeal cavity (Delattre & Freeman, 1968), with the latter

accompanied by lateral bracing and posterior tongue body lowering. Viewing the LPC spectrum while simultaneously shaping the child’s production via articulator placement cues can help determine which combination of elements to focus on in treatment and scaffold the accuracy of novel tongue movements that differ from the existing habituated, yet off-target, motor plan. For example, some speakers produce /l/ with pharyngeal constriction, so shaping from /l/ may scaffold production of a necessary element for /ɻ/ (Shriberg, 1975). The clinician can suggest an articulator placement cue with a direct reference to the LPC spectrum (e.g., “Try moving your tongue back and watch what happens to the wave”). For a detailed list of articulator placement cues and shaping strategies for American English /ɻ/, see the work of Preston et al. (2020). Although Preston et al. (2020) focuses on traditional (i.e., nonbiofeedback) treatment, clinicians are encouraged to incorporate the same strategies as articulator placement cues during delivery of VAB.

Figure 8. Comparison of correct and incorrect /ɻ/ productions.



Goal Selection

The American English rhotic phoneme /ɻ/ can occur in different positions in the syllable, including prevocalic position as in *red*, syllable nucleus as in *her*, and post-vocalic position (sometimes described as the offglide of a rhotic diphthong) as in *deer* and *door*. These positional variants have slightly different acoustic and articulatory characteristics (McGowan et al., 2004), and identifying which one to target first is a clinically important question. We suggest initially selecting a stressed syllabic /ɻ/ for several reasons. From a developmental perspective, there is reason to believe that vocalic targets emerge earlier in

typically developing speakers, which suggests that vocalic /ɜ/ may also be likely to emerge in treatment before other variants (Klein et al., 2013; McGowan et al., 2004). Additionally, syllabic /ɜ/ has been shown to have a longer duration relative to /ɪ/ in a syllable onset position (e.g., *road, tray*). This longer duration generally makes it easier for a client to identify target formant peaks using the LPC spectrum. However, it is also possible that a client may demonstrate an optimal response to /ɪ/ in postvocalic position (e.g., *care, fear, and car*), particularly when the off-glide is influenced in a facilitative way by the phonetic context of the preceding vowel. For example, the clients who struggle to form a constriction in the pharyngeal region when attempting a rhotic production may have more success in postvocalic context following a low back vowel (e.g., /ɑ/) due to the facilitative nature of the neighboring articulatory context (Boyce, 2015), which could justify selecting “are” as an early target as well. Therefore, clinicians are encouraged to carefully assess several contexts that may be facilitative for each learner.

Treatment Sessions

Suggested Outline of a Visual–Acoustic Treatment Session Using LPC Spectra

Prepractice. Treatment sessions are usually initiated following one to two introductory sessions explaining production of /ɪ/ and the connection to the LPC spectrum, as outlined previously. Each session begins with a relatively unstructured prepractice period in which clients are encouraged to explore a variety of articulatory combinations to try to align their spectrum with a preselected template. Initially, guidance to explore the lingual and labial movement may be offered without reference to specific articulators (e.g., “While you are looking at the image on the screen <clinician points to spectral peaks on the computer screen>, try to move your tongue around in your mouth to make the peaks or bumps change. It may not sound like an /ɪ/ right away, but that’s okay for now”). Clinicians provide general encouragement and then systematically begin to suggest specific articulatory cues to facilitate acoustic and perceptual changes. Introduction of articulator cues may be spaced out over time (e.g., limited to one type of articulatory cue used throughout the duration of a session) in order to limit cognitive load. Alternatively, cues may be grouped together within a session. Next, the clinician and the client review the dynamic LPC image to determine the impact of a given cue on the spectral shape. Additional articulatory cues (e.g., raise the tongue blade, lower the tongue dorsum, and retract the tongue root) can be added as the child becomes more comfortable understanding formant peaks. The articulatory cues found to facilitate the most change during prepractice can form the focus of subsequent structured practice trials.

The prepractice period involves relatively unstructured, highly interactive elicitation of targets to help the client begin to connect movements of the articulators with changes in the acoustic waveform. In the early acquisition phase of learning, we recommend that prepractice comprise a relatively large percentage of the total session time (e.g., roughly 50% or 20–25 min in a 50-min session) and elicit a limited range of /ɪ/ targets. In this early phase, prepractice may include several different syllables containing /ɪ/, selected depending on the individualized needs of each client. In our current research studies (e.g., McAllister et al., 2020), we select one stimulus item from each of the five target /ɪ/ variants, beginning with elicitation of /ɜ/ (see *Goal Selection* for more details). The clinician should work to elicit each target using verbal models, VAB, and articulator placement cues. In our research, we advance to the next target within prepractice when the client has produced the target 3 times in a fully correct fashion or completed 10 unsuccessful attempts, whichever comes first. Prepractice may be terminated after (a) a set number of trials or time duration or (b) a set number of correct productions of the selected target variants. As treatment advances, the prepractice duration may decrease if the client accurately produces the selected target variants at least 3 times; for instance, the duration could be reduced from approximately 50% to 10% (5 min of a 50-min session). The /ɪ/ variants selected in prepractice are typically carried over into the period of structured practice that follows.

Structured practice. A period of high-intensity structured practice typically follows the prepractice phase of treatment. Previous literature investigating the efficacy of VAB intervention for residual /ɪ/ distortions has varied in the intensity of treatment with respect to session frequency, session duration, and number of trials elicited per session (see Table 2). Given the recent evidence of the importance of dosage in biofeedback treatment (Hitchcock et al., 2019; Peterson et al., 2022), we recommend targeting 100–150 trials at a minimum.

One challenge widely acknowledged in the context of biofeedback treatment (e.g., Gibbon & Paterson, 2006; McAllister Byun & Hitchcock, 2012) is the possibility that new skills learned in treatment may not generalize to a context in which biofeedback is not available. Once the client establishes an accurate /ɪ/ production, we recommend structuring practice to maximize generalization of the newly acquired sound. In particular, we recommend that the main duration of the session should emphasize production of the target rhotic contexts at a stage of complexity where accuracy is achievable but not too difficult or too easy, also known as an optimal *challenge point* level (Guadagnoli & Lee, 2004). This can be operationalized as the level where the client can achieve 50%–80% accuracy within the treatment setting. In our clinical research, we offer adaptive difficulty in structured practice using the

Challenge Point Program (CPP; McAllister et al., 2021), a free and open source PC-based software that encodes a structured version of a challenge point hierarchy (Matthews et al., 2021; Rvachew & Brosseau-Lapre, 2012) for /i/ practice.⁴ The CPP was designed to make it feasible for clinicians to elicit multiple treatment trials while adaptively increasing or decreasing task difficulty based on within-session performance. The adaptive behavior of the CPP is determined by three within-session parameters adjusted on a rotating schedule (see Appendix B). The parameters alter the functional task difficulty by changing the frequency with which biofeedback is made available, the mode of elicitation (e.g., imitation vs. independent reading), and the complexity of target productions presented (e.g., syllables, words, and phrases) based on the participant's accuracy over 10 trials. If accuracy is 80% or better, the CPP adjusts one parameter to increase difficulty in the next block. If accuracy again reaches or exceeds 80%, another manipulation is added to further increase difficulty. If accuracy falls at or below 50%, these manipulations are withdrawn in reverse order of application to reduce difficulty. As a result, biofeedback is faded and the production task becomes progressively more challenging as client accuracy improves. Further detail on the nature of CPP can be found in the work of McAllister et al. (2021).

Another significant consideration for clinicians adopting VAB comes from evidence that biofeedback strategies may be most effective during the early acquisition stages of speech learning (Fletcher et al., 1991; Gibbon & Paterson, 2006; Preston et al., 2018, 2019; Volin, 1998). The premise that VAB may be most effective for establishing new motor patterns is consistent with the broader body of research investigating principles of motor skill learning. Certain parameters of practice have proved more facilitative of the initial acquisition of a motor plan, whereas other practice conditions may maximize retention and transfer (see review in Maas et al., 2008). Qualitative *knowledge of performance* feedback, defined in this context as feedback that helps the speaker identify how a sound was produced, seems to offer the greatest advantage for the client when the motor task is novel or the nature of the target is unclear to the client (Newell et al., 1990). In later phases of learning, detailed knowledge of performance feedback has been reported to be less effective and potentially even detrimental to learning. On the other hand, *knowledge of results* feedback, defined as identifying if a sound was produced correctly or incorrectly, has been shown to be most effective in later phases of learning (Maas et al., 2008). Regardless, a critical element of feedback is that it is always based on the clinician's judgment of the accuracy of how the production sounds, which should be prioritized over the "correctness" of

the acoustic display image. That is, VAB is a tool for generating a perceptually correct sounding /i/. Matching a template is not explicitly the goal if it does not result in a production that sounds correct. An expert clinician's perception of the child's attempt should determine the accuracy of the production, which, ideally, is supported by the visual display.

Emphasizing the identification of formant peak changes secondary to articulatory alterations during sound acquisition can be considered knowledge of performance feedback, given the correlation between the behavior (articulatory movements) and the resulting visual change (altered formant peaks). Therefore, the principles of motor skill learning suggest that biofeedback is likely to be most effective in the earliest stages of learning, and its utility may decline over time as the learner becomes more proficient (McAllister Byun & Campbell, 2016; Peterson et al., 2022; Preston et al., 2018). In keeping with this theoretical framing, two separate studies including both traditional and biofeedback treatment in similar counterbalanced study designs reported a measurable advantage when biofeedback was provided prior to traditional treatment (McAllister Byun & Campbell, 2016; Preston et al., 2019).

Additional Considerations to Optimize Clinical Outcomes

Computer Monitor Sight Line

Several additional factors may influence the efficacy of biofeedback treatment. For example, it is crucial for clients to view the screen image during attempts to modify their articulatory behaviors. While this may seem obvious, in our experience, many clients tend to look toward the clinician when being given directions. As a result, they are not watching the screen for related changes in the dynamic image. This may limit the client's ability to connect articulatory movements with formant changes reflected in the LPC spectrum, necessitating redirection to focus on the screen. Frequent redirection may be necessary to maintain focus on the visual display, particularly for younger clients or clients with comorbid diagnoses such as attention-deficit/hyperactivity disorder. Because repetitive reminders may increase client frustration, we suggest framing the learning task so the child has an active role in monitoring the visual output (e.g., after each trial, the child describes what they saw on the visual display with specific reference to the movement of the "bumps").

Client Posture

We also recommend that the client is seated in a fully upright position when viewing the screen image. A lowered chin typically limits the range of articulatory movement, potentially limiting the child's ability to alter existing articulatory behaviors. Additionally, slouching limits thoracic volume for respiratory support during phonation. This can

⁴The CPP is available at <http://blog.umd.edu/cpp/download/>.

result in a low-intensity or distorted acoustic signal, which may result in limited or absent peaks on the LPC spectrum. Raising the table height or the computer monitor may mitigate these situations and optimize learning opportunities.

Microphone

In order to maximize the accuracy of the LPC spectrum representing the client's /s/ production, it is important to achieve a good signal-to-noise ratio. In this scenario, the signal is the client's voice compared with the level of background noise. The strength of this input signal will be influenced by the device used for voice recording. While it is possible to use the sound card on a computer, we recommend using a dedicated external microphone for the best signal-to-noise ratio. There are several types of directional patterns for microphones, including omnidirectional and unidirectional. In general, a unidirectional microphone is preferable to an omnidirectional microphone, because it is less likely to pick up unwanted background sound. However, when using a unidirectional microphone, it is important to remain attentive to the distance and angle between the client's mouth and the microphone. The optimal mouth-to-microphone distance may vary across speakers and devices, but we recommend testing different distances until a clear signal is achieved and then encouraging the client to remain at that distance as consistently as possible.

Other Benefits of VAB

VAB offers several unique benefits compared with other forms of visual biofeedback. It is the least invasive form of biofeedback, requiring only a microphone and a computer screen. Other visualization technologies used to facilitate acquisition of speech targets require direct contact with the speech structures (for instance, ultrasound requires direct skin contact via a probe held beneath the chin, and electropalatography requires intraoral placement of the pseudopalate used to register and display areas of linguopalatal contact; Bernhardt et al., 2003; Dagenais et al., 1994; Hitchcock et al., 2017; McAllister Byun et al., 2014; Preston et al., 2013). VAB also tends to be the least expensive biofeedback technique. As noted previously, some software for acoustic analysis and speech visualization programs can be downloaded at little or no cost (e.g., staRt). Clinical software programs such as Sona-Speech (PENTAX Medical) are not free but are typically less costly than the hardware required for articulatory types of biofeedback such as ultrasound. Last, the lack of additional hardware other than an external microphone also renders VAB more amenable to delivery via telepractice compared with other visual biofeedback techniques, thus increasing the treatment delivery options available via telepractice in speech pathology. For further discussion of the use of VAB via telepractice, see the work of Peterson et al. (2022).

It is sometimes suggested that VAB may be harder for clients and clinicians to interpret because formant patterns are more abstract than a direct display of articulator shape or contacts. However, there is a lack of evidence directly investigating this question of ease of interpretation across biofeedback types. As noted above, the only published evidence directly comparing the two types of biofeedback (VAB and ultrasound) reported one participant with a slight advantage for VAB over ultrasound, with the remaining six participants showing no significant difference between conditions (Benway et al., 2021).

Conclusions

The use of visual biofeedback to treat individuals with speech sound errors who show a limited response to traditional interventions has grown significantly over the past 30 years. The authors' collective findings demonstrate that VAB can facilitate perceptually and acoustically correct rhotic production in children with whose residual distortions of /s/ have not responded to traditional methods of intervention (e.g., McAllister Byun & Hitchcock, 2012; Peterson et al., 2022). The largely successful outcomes of these studies, as well as the evidence of successful use of VAB with L2 learners (e.g., Li et al., 2019) and individuals with hearing loss (e.g., Ertmer & Maki, 2000), demonstrate the potential for spectral/spectrographic displays to help learners effectively alter their own speech production patterns.

The suggested strategies for using VAB proposed in this tutorial are based on approximately 10 years of research experience by the authors. Our suggested course of treatment includes a clear introductory phase to familiarize the client with the technology, an acquisition phase marked by intensive VAB use, and a generalization phase in which the use of VAB is faded while target complexity is simultaneously increased. While not all of the strategies recommended here have been the subject of rigorous experimental manipulations, each has been refined over the course of numerous research studies. In summary, the strategies offered here may serve as a guideline for clinicians planning to incorporate the use of visual-acoustic technology into their clinical toolbox. Increased clinical adoption of VAB may facilitate improved outcomes for clients with a range of different speech goals.

Author Contributions

Elaine R. Hitchcock: Conceptualization (Lead), Writing – original draft (Lead), Writing – review & editing (Lead), Visualization (Equal). **Laura C. Ochs:** Conceptualization (Equal), Writing – original draft (Equal), Writing – review & editing (Equal), Visualization (Equal).

Michelle T. Swartz: Project administration (Supporting), Writing – original draft, Writing – review & editing (Supporting). **Megan C. Leece:** Project administration (Supporting), Resources (Equal), Visualization (Equal), Writing – review & editing (Supporting). **Jonathan L. Preston:** Project administration (Equal), Writing – original draft (Equal), Writing – review & editing (Equal). **Tara McAllister:** Funding acquisition (Lead), Project administration (Lead), Writing – original draft (Equal), Writing – review & editing (Equal).

Data Availability Statement

All data generated or analyzed during this study are included in this published article and its supplemental material files.

Acknowledgments

Research reported in this publication was supported by National Institute on Deafness and Other Communication Disorders Grant R15DC019775 (principal investigator: E. Hitchcock) and R01DC017476 (principal investigator: T. McAllister). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The authors gratefully acknowledge Roberta Lazarus, Sarah Granquist, Nina Benway, Lynne Harris, and Samantha Ayala for their vital contributions as clinical partners in this study.

References

- Akahane-Yamada, R., McDermott, E., Adachi, T., Kawahara, H., & Pruitt, J. S. (1998). *Computer-based second language production training by using spectrographic representation and HMM-based speech recognition scores*. Paper presented at the Fifth International Conference on Spoken Language Processing.
- Awan, S. N. (2013). *Applied speech & voice analysis using the KayPENTAX speech product line*. PENTAXMedical.
- 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(1), 36–45. <https://doi.org/10.1080/14417040601101037>
- Beeson, P. M., & Robey, R. R. (2006). Evaluating single-subject treatment research: Lessons learned from the aphasia literature. *Neuropsychology Review*, 16, 161–169. <https://doi.org/10.1007/s11065-006-9013-7>
- Benway, N. R., Hitchcock, E. R., McAllister, T., Feeny, G. T., Hill, J., & Preston, J. L. (2021). Comparing biofeedback types for children with residual /l/ errors in American English: A single-case randomization design. *American Journal of Speech-Language Pathology*, 30(4), 1819–1845. https://doi.org/10.1044/2021_AJSLP-20-00216
- Bernhardt, B., Gick, B., Bacsfalvi, P., & Adler-Bock, M. (2005). Ultrasound in speech therapy with adolescents and adults. *Clinical Linguistics & Phonetics*, 19(6–7), 605–617. <https://doi.org/10.1080/02699200500114028>
- Bernhardt, B., 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(3), 199–216. <https://doi.org/10.1080/0269920031000071451>
- BITS Lab NYU. (2020). staRt (Vers. 3.4.8). [Mobile application software]. <https://apps.apple.com/us/app/bits-lab-start/id1198658004>
- Boyce, S., & Espy-Wilson, C. Y. (1997). Coarticulatory stability in American English /r/. *The Journal of the Acoustical Society of America*, 101(6), 3741–3753. <https://doi.org/10.1121/1.418333>
- Boyce, S. E. (2015). The articulatory phonetics of /r/ for residual speech errors. *Seminars in Speech and Language*, 36(4), 257–270. <https://doi.org/10.1055/s-0035-1562909>
- Brady, K. W., Duewer, N., & King, A. M. (2016). The effectiveness of a multimodal vowel-targeted intervention in accent modification. *Contemporary Issues in Communication Science and Disorders*, 43, 23–34. https://doi.org/10.1044/cicsd_43_S_23
- Busk, P. L., & Serlin, R. C. (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). Erlbaum.
- Campbell, H. M., Harel, D., & McAllister Byun, T. (2017). Selecting an acoustic correlate for automated measurement of /r/ production in children. *The Journal of the Acoustical Society of America*, 141(5), 3572–3572. <https://doi.org/10.1121/1.4987592>
- Carey, M. (2004). CALL visual feedback for pronunciation of vowels. *CALICO Journal*, 21(3), 571–601. <https://doi.org/10.1558/cj.v21i3.571-601>
- Chiba, T., & Kajiyama, M. (1941). *The vowel: Its nature and structure*. Tokyo-Kaiseikan Pub. Co., Ltd.
- Cialdella, L., Kabakoff, H., Preston, J., Dugan, S., Spencer, C., Boyce, S., Whalen, D., & McAllister, T. (2021). Auditory-perceptual acuity in rhotic misarticulation: Baseline characteristics and treatment response. *Clinical Linguistics & Phonetics*, 35(1), 19–42. <https://doi.org/10.1080/02699206.2020.1739749>
- Crawford, E. (2007). *Acoustic signals as visual biofeedback in the speech training of hearing impaired children*. University of Canterbury. Communication Disorders.
- Culbertson, D. N., & Kricos, P. B. (2002). Language and speech of the deaf and hard of hearing. In R. L. Schow & M. A. Ling (Eds.), *Introduction to Audiologic rehabilitation* (4th ed., pp. 183–224). Allyn & Bacon.
- Dagenais, P. A., Critz-Crosby, P., & Adams, J. B. (1994). Defining and remediating persistent lateral lisps in children using Electropalatography: Preliminary findings. *American Journal of Speech-Language Pathology*, 3(3), 67–76. <https://doi.org/10.1044/1058-0360.0303.67>
- Delattre, P., & Freeman, D. C. (1968). A dialect study of American r's by x-ray motion picture. *Linguistics*, 6(44), 29–68. <https://doi.org/10.1515/ling.1968.6.44.29>
- Dowd, A., Smith, J., & Wolfe, J. (1998). Learning to pronounce vowel sounds in a foreign language using acoustic measurements of the vocal tract as feedback in real time. *Language and Speech*, 41(1), 1–20. <https://doi.org/10.1177/002383099804100101>
- Ertmer, D. J., & Maki, J. E. (2000). A comparison of speech training methods with deaf adolescents: Spectrographic versus noninstrumental instruction. *Journal of Speech and Hearing Research*, 43(6), 1509–1523. <https://doi.org/10.1044/jslhr.4306.1509>
- Ertmer, D. J., Stark, R. E., & Karlan, G. R. (1996). Real-time spectrographic displays in vowel production training with

- children who have profound hearing loss. *American Journal of Speech-Language Pathology*, 5(4), 4–16. <https://doi.org/10.1044/1058-0360.0504.04>
- Fant, G.** (1960). *Acoustic theory of speech production*. Mouton & Co.
- 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(4), 929–943. <https://doi.org/10.1044/jshr.3404.929>
- Freedman, S. E., Maas, E., Caligiuri, M. P., Wulf, G., & Robin, D. A.** (2007). Internal versus external: Oral-motor performance as a function of attentional focus. *Journal of Speech, Language, and Hearing Research*, 50(1), 131–136. [https://doi.org/10.1044/1092-4388\(2007/011\)](https://doi.org/10.1044/1092-4388(2007/011))
- Gibbon, F., Dent, H., & Hardcastle, W.** (1993). Diagnosis and therapy of abnormal alveolar stops in a speech disordered child using electropalatography. *Clinical Linguistics & Phonetics*, 7(4), 247–267. <https://doi.org/10.1080/02699209308985565>
- 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(3), 275–292. <https://doi.org/10.1191/0265659006ct308xx>
- Gick, B., Bacsfalvi, P., Bernhardt, B. M., Oh, S., Stolar, S., & Wilson, I.** (2007). A motor differentiation model for liquid substitutions in children's speech. In *Proceedings of meetings on acoustics 153ASA* (Vol. 1, No. 1, p. 060003). Acoustical Society of America.
- Guadagnoli, M. A., & Lee, T. D.** (2004). Challenge point: A framework for conceptualizing the effects of various practice conditions in motor learning. *Journal of Motor Behavior*, 36(2), 212–224. <https://doi.org/10.3200/JMBR.36.2.212-224>
- Guenther, F. H., Hampson, M., & Johnson, D.** (1998). A theoretical investigation of reference frames for the planning of speech movements. *Psychological Review*, 105(4), 611–633. <https://doi.org/10.1037/0033-295X.105.4.611-633>
- Hitchcock, E. R., Cabbage, K. L., Swartz, M. T., & Carrell, T. D.** (2020). Measuring speech perception using the wide-range acoustic accuracy scale: Preliminary findings. *Perspectives of the ASHA Special Interest Groups*, 5(4), 1098–1112. https://doi.org/10.1044/2020_PERSP-20-00037
- Hitchcock, E. R., McAllister Byun, T., Swartz, M., & Lazarus, R.** (2017). Efficacy of electropalatography for treating misarticulation of /r/. *American Journal of Speech-Language Pathology*, 26(4), 1141–1158. https://doi.org/10.1044/2017_AJSLP-16-0122
- Hitchcock, E. R., Swartz, M. T., & Lopez, M.** (2019). Speech sound disorder and visual biofeedback intervention: A preliminary investigation of treatment intensity. *Seminars in Speech and Language*, 40(2), 124–137. <https://doi.org/10.1055/s-0039-1677763>
- Kartushina, N., Hervais-Adelman, A., Frauenfelder, U. H., & Golestani, N.** (2015). The effect of phonetic production training with visual feedback on the perception and production of foreign speech sounds. *The Journal of the Acoustical Society of America*, 138(2), 817–832. <https://doi.org/10.1121/1.4926561>
- Kartushina, N., Hervais-Adelman, A., Frauenfelder, U. H., & Golestani, N.** (2016). Mutual influences between native and non-native vowels in production: Evidence from short-term visual articulatory feedback training. *Journal of Phonetics*, 57, 21–39. <https://doi.org/10.1016/j.wocn.2016.05.001>
- King, H., & Ferragne, E.** (2020). Loose lips and tongue tips: The central role of the /r/-typical labial gesture in Anglo-English. *Journal of Phonetics*, 80, 100978. <https://doi.org/10.1016/j.wocn.2020.100978>
- Klein, H. B., McAllister Byun, T., Davidson, L., & Grigos, M. I.** (2013). A multidimensional investigation of children's /r/ productions: Perceptual, ultrasound, and acoustic measures. *American Journal of Speech-Language Pathology*, 22(3), 540–553. [https://doi.org/10.1044/1058-0360\(2013/12-0137\)](https://doi.org/10.1044/1058-0360(2013/12-0137))
- Ladefoged, P.** (1996). *Elements of acoustic phonetics*. University of Chicago Press. <https://doi.org/10.7208/chicago/9780226191010.001.0001>
- Lee, S., Potamianos, A., & Narayanan, S.** (1999). Acoustics of children's speech: Developmental changes of temporal and spectral parameters. *The Journal of the Acoustical Society of America*, 105(3), 1455–1468. <https://doi.org/10.1121/1.426686>
- Li, J. J., Ayala, S., Harel, D., Shiller, D. M., & McAllister, T.** (2019). Individual predictors of response to biofeedback training for second-language production. *The Journal of the Acoustical Society of America*, 146(6), 4625–4643. <https://doi.org/10.1121/1.5139423>
- Maas, E., Robin, D. A., Ausermann Hula, S. N., Freedman, S. E., Wulf, G., Ballard, K. J., & Schmidt, R. A.** (2008). Principles of motor learning in treatment of motor speech disorders. *American Journal of Speech-Language Pathology*, 17(3), 277–298. [https://doi.org/10.1044/1058-0360\(2008/025\)](https://doi.org/10.1044/1058-0360(2008/025))
- Maki, J. E., & Streff, M. M.** (1978). *Clinical evaluation of the speech spectrographic display with hearing impaired adults*. Paper presented at the American Speech and Hearing Association.
- Matthews, T., Barbeau-Morrison, A., & Rvachew, S.** (2021). Application of the challenge point framework during treatment of speech sound disorders. *Journal of Speech, Language, and Hearing Research*, 64(10), 3769–3785. https://doi.org/10.1044/2021_JSLHR-20-00437
- McAllister, T., Preston, J. L., Hitchcock, E. R., & Hill J.** (2020). Protocol for Correcting Residual Errors with Spectral, ULtrasound, Traditional Speech therapy Randomized Controlled Trial (C-RESULTS RCT). *BMC Pediatrics*, 20(1), 66. <https://doi.org/10.1186/s12887-020-1941-5>
- McAllister Byun, T.** (2017). Efficacy of visual-acoustic biofeedback intervention for residual rhotic errors: A single-subject randomization study. *Journal of Speech, Language, and Hearing Research*, 60(5), 1175–1193. https://doi.org/10.1044/2016_JSLHR-S-16-0038
- McAllister Byun, T., & Campbell, H.** (2016). Differential effects of visual-acoustic biofeedback intervention for residual speech errors. *Frontiers in Human Neuroscience*, 10, 567. <https://doi.org/10.3389/fnhum.2016.00567>
- McAllister Byun, T., Campbell, H., Carey, H., Liang, W., Park, T. H., & Svirsky, M.** (2017). Enhancing intervention for residual rhotic errors via app-delivered biofeedback: A case study. *Journal of Speech, Language, and Hearing Research*, 60(6S), 1810–1817. https://doi.org/10.1044/2017_JSLHR-S-16-0248
- 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(3), 207–221. [https://doi.org/10.1044/1058-0360\(2012/11-0083\)](https://doi.org/10.1044/1058-0360(2012/11-0083))
- McAllister Byun, T., Hitchcock, E. R., & Swartz, M. T.** (2014). Retroflex versus bunched in treatment for rhotic misarticulation: Evidence from ultrasound biofeedback intervention. *Journal of Speech, Language, and Hearing Research*, 57(6), 2116–2130. https://doi.org/10.1044/2014_JSLHR-S-14-0034
- McAllister Byun, T., Swartz, M. T., Halpin, P. F., Szeredi, D., & Maas, E.** (2016). Direction of attentional focus in biofeedback treatment for /r/ misarticulation. *International Journal of Language & Communication Disorders*, 51(4), 384–401. <https://doi.org/10.1111/1460-6984.12215>
- McAllister, T., Hitchcock, E. R., & Ortiz, J.** (2021). Computer-assisted challenge point intervention for residual speech errors. *Perspectives of the ASHA Special Interest Groups*, 6(1), 214–229. https://doi.org/10.1044/2020_PERSP-20-00191

- McGowan, R. S., Nittrouer, S., & Manning, C. J. (2004). Development of [ɹ] in young, midwestern, American children. *The Journal of the Acoustical Society of America*, 115(2), 871–884. <https://doi.org/10.1121/1.1642624>
- Mielke, J., Baker, A., & Archangeli, D. (2016). Individual-level contact limits phonological complexity: Evidence from bunched and retroflex /ɹ/. *Language*, 92(1), 101–140. <https://doi.org/10.1353/lan.2016.0019>
- 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(4), 536–552. <https://doi.org/10.1080/00222895.1990.10735527>
- Olson, D. J. (2014). Benefits of visual feedback on segmental production in the L2 classroom. *Language Learning & Technology*, 18(3), 173–192. <https://doi.org/10.10125/44389>
- PENTAX Medical. (2019). *Computerized Speech Lab (CSL), Model 4500*. [software]. <https://www.pentaxmedical.com/pentax/en/99/1/Computerized-Speech-Lab-CSL>
- Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of the vowels. *The Journal of the Acoustical Society of America*, 24(2), 175–184. <https://doi.org/10.1121/1.1906875>
- Peterson, L., Savarese, C., Campbell, T., Ma, Z., Simpson, K. O., & McAllister, T. (2022). Telepractice treatment of residual rhotic errors using app-based biofeedback: A pilot study. *Language, Speech, and Hearing Services in Schools*, 256–274. https://doi.org/10.1044/2021_LSHSS-21-00084
- Preston, J. L., Benway, N. R., Leece, M. C., Hitchcock, E. R., & McAllister, T. (2020). Tutorial: Motor-based treatment strategies for /r/ distortions. *Language, Speech, and Hearing Services in Schools*, 51(4), 966–980. https://doi.org/10.1044/2020_LSHSS-20-00012
- Preston, J. L., Brick, N., & Landi, N. (2013). Ultrasound biofeedback treatment for persisting childhood apraxia of speech. *American Journal of Speech-Language Pathology*, 22(4), 627–643. [https://doi.org/10.1044/1058-0360\(2013\)12-0139](https://doi.org/10.1044/1058-0360(2013)12-0139)
- Preston, J. L., McAllister, T., Phillips, E., Boyce, S., Tiede, M., Kim, J. S., & Whalen, D. H. (2018). Treatment for residual rhotic errors with high- and low-frequency ultrasound visual feedback: A single-case experimental design. *Journal of Speech, Language, and Hearing Research*, 61(8), 1875–1892. https://doi.org/10.1044/2018_JSLHR-S-17-0441
- Preston, J. L., McAllister, T., Phillips, E., Boyce, S., Tiede, M., Kim, J. S., & Whalen, D. H. (2019). Remediating residual rhotic errors with traditional and ultrasound-enhanced treatment: A single-case experimental study. *American Journal of Speech-Language Pathology*, 28(3), 1167–1183. https://doi.org/10.1044/2019_AJSLP-18-0261
- Preston, J. L., McCabe, P., Rivera-Campos, A., Whittle, J. L., Landry, E., & Maas, E. (2014). Ultrasound visual feedback treatment and practice variability for residual speech sound errors. *Journal of Speech, Language, and Hearing Research*, 57(6), 2102–2115. https://doi.org/10.1044/2014_JSLHR-S-14-0031
- Ruscello, D. M. (1995). Visual feedback in treatment of residual phonological disorders. *Journal of Communication Disorders*, 28(4), 279–302. [https://doi.org/10.1016/0021-9924\(95\)00058-x](https://doi.org/10.1016/0021-9924(95)00058-x)
- Rvachew, S., & Brosseau-Lapre, F. (2012). An input-focused intervention for children with developmental phonological disorders. *SIG 1 Perspectives on Language Learning and Education*, 19(1), 31–35. <https://doi.org/10.1044/lle19.1.31>
- Schmidt, A. M. (2007). Evaluating a new clinical palatometry system. *Advances in Speech Language Pathology*, 9(1), 73–81. <https://doi.org/10.1080/14417040601123650>
- Shriberg, L. D. (1975). A response evocation program for /ɹ/. *Journal of Speech and Hearing Disorders*, 40(1), 92–105. <https://doi.org/10.1044/jshd.4001.92>
- Shriberg, L. D., Flipsen, P., Jr., Karlsson, H. B., & McSweeney, J. L. (2001). Acoustic phenotypes for speech-genetics studies: An acoustic marker for residual /ɹ/ distortions. *Clinical Linguistics & Phonetics*, 15(8), 631–650. <https://doi.org/10.1080/02699200110069429>
- Shuster, L. I. (1998). The perception of correctly and incorrectly produced /r/. *Journal of Speech, Language, and Hearing Research*, 41(4), 941–950. <https://doi.org/10.1044/jslhr.4104.941>
- Shuster, L. I., Ruscello, D. M., & Smith, K. D. (1992). Evoking [r] using visual feedback. *American Journal of Speech-Language Pathology*, 1(3), 29–34. <https://doi.org/10.1044/1058-0360.0103.29>
- 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(2), 37–44. <https://doi.org/10.1044/1058-0360.0402.37>
- Stark, R. E. (1971). The use of real-time visual displays of speech in the training of a profoundly deaf non-speaking child: A case report. *Journal of Speech and Hearing Disorders*, 36(3), 397–409. <https://doi.org/10.1044/jshd.3603.397>
- Stavness, I., Gick, B., Derrick, D., & Fels, S. (2012). Biomechanical modeling of English /r/ variants. *The Journal of the Acoustical Society of America*, 131(5), EL355–EL360. <https://doi.org/10.1121/1.3695407>
- Sugden, E., Lloyd, S., Lam, J., & Cleland, J. (2019). Systematic review of ultrasound visual biofeedback in intervention for speech sound disorders. *International Journal of Language & Communication Disorders*, 54(5), 705–728. <https://doi.org/10.1111/1460-6984.12478>
- Swartz, M. T., Hitchcock, E. R., & Boyle, M. (2018). *Improving prosodic variation in a patient with primary progressive apraxia of speech using visual-acoustic biofeedback*. American Speech-Language-Hearing Association (ASHA) Convention.
- Tiede, M. K., Boyce, S. E., Holland, C. K., & Choe, K. A. (2004). A new taxonomy of American English /r/ using MRI and ultrasound. *The Journal of the Acoustical Society of America*, 115(5), 2633–2634. <https://doi.org/10.1121/1.4784878>
- Utianski, R. L., Clark, H. M., Duffy, J. R., Botha, H., Whitwell, J. L., & Josephs, K. A. (2020). Communication limitations in patients with progressive apraxia of speech and aphasia. *American Journal of Speech-Language Pathology*, 29(4), 1976–1986. https://doi.org/10.1044/2020_AJSLP-20-00012
- 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(1), 81–90. <https://doi.org/10.1044/1058-0360.0701.81>
- Zemlin, W. R. (1998). *Speech and hearing science: Anatomy and physiology* (4th ed.). Allyn & Bacon.
- Zhou, X., Espy-Wilson, C. Y., Boyce, S., Tiede, M., Holland, C., & Choe, A. (2008). A magnetic resonance imaging-based articulatory and acoustic study of ‘retroflex’ and ‘bunched’ American English /r/. *The Journal of the Acoustical Society of America*, 123(6), 4466–4481. <https://doi.org/10.1121/1.2902168>

Appendix A

Introduction to Visual–Acoustic Biofeedback With Vowels

Introduction to VAB with vowels

Main point(s)

Script

Show how to use visual display to make a better /i/ sound.

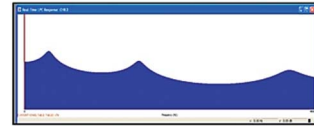
“We’re going to use a computer program that makes a picture of your speech sounds. We’re going to use this to try and help you say a better /i/ sound. Before we use the computer, we need to learn a little bit about what the computer image is telling us, so we know what to look for on the screen.”

Different sounds have different shapes on the LPC spectrum.

“Here’s what the pictures will look like. The blue wave is actually a picture of your speech. Different sounds will have **bumps/peaks*** in the wave in different places. As you make different speech sounds, you’ll see this wave move around.”

<Show image/live model of LPC Spectrum.>

LPC Spectrum



Note: The clinician is free to choose the term “bump” or “peak” based on the client’s age and/or general preference.

Moving tongue around in the mouth changes the tongue shape which moves the blue peaks around on the image.

“Let’s see what it looks like. Watch how the wave moves around when I talk.”

<Demonstrate moving wave as you talk.>

“Now you try making the wave move around. When we use the computer program we need you to sit up straight and speak directly into the microphone while looking at the screen.”

<Have client move wave while speaking.>

“Now watch this—I can make the wave do different things when I say different vowel sounds.

The peaks/bumps will be in different places for different sounds, like /i/ and /a/.”

<Demonstrate movement of wave when moving from /i/ to /a/>

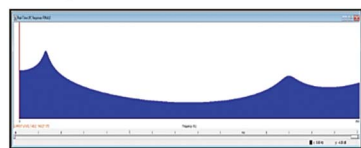
<Sustain the following sounds for 2–3 s each: /i, a, u/>

< Have client sustain vowels for 2–3 s.>

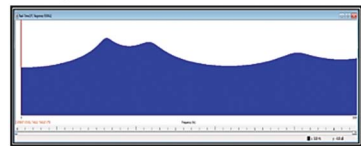
“Now watch this—I can make a tracing of the sound I made.”

<Demonstrate how a template can be created of /i/ using your own vowel production.>

LPC spectrum of /i/ vowel

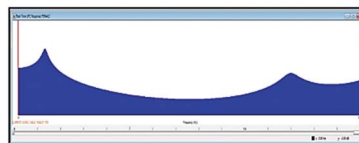


LPC spectrum of /a/ vowel

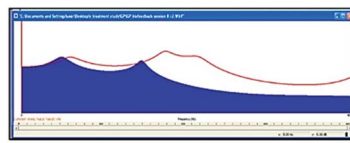


Make clinician template of /i/ and demonstrate matching the template.

LPC spectrum of /i/ vowel



LPC spectrum with template



“Look at this red line. It’s like a tracing or the outline of the sound I just made. Whenever I make that same sound again, the bumps/peaks on the blue wave should line up with the bumps/peaks on that red line. It won’t always line up perfectly, because every sound is different, but it should be just about the same.” <Sustain an /i/ vowel.> “Do you see how the bumps line up?”

Load preselected speaker template match for /a/, /i/, /u/, and encourage child to explore/match template shape.

“Now I want you to try. But our voices are different, right? So, if you try to match my red line, it may not line up too well. But we also have some lines that we traced from kids whose voices sound more like yours. We’ll have you try to match one of those lines instead.” <Select the preselected speaker template match for the /a/ vowel.> “This line shows the shape of another kid’s /a/ sound. I want you to say your best /a/ sound into the microphone and see if you can make the bumps/peaks line up.” <Allow the client to play around until a satisfactory match with the template is reached. If client can’t achieve a match, select a different template.> <Load the /i/, /u/ templates from the same preselected speaker templates and encourage child to move tongue to match each template.>

Appendix B

Within-Session Levels in the Challenge Point Program (CPP) Software (McAllister et al., 2021)

Level	Biofeedback frequency	Mode of elicitation	Stimulus complexity
1	100%	Imitate clinician's model	1 syllable simple
2	50%	Imitate clinician's model	1 syllable simple
3	50%	Read independently	1 syllable simple
4	50%	Read independently	1 syllable with competing /l/ or /w/
5	0%	Read independently	1 syllable with competing /l/ or /w/
6	0%	Imitation with prosodic manipulation	1 syllable with competing /l/ or /w/
7	0%	Imitation with prosodic manipulation	2 syllables simple
8	0%	Independent reading with prosodic manipulation	2 syllables simple
9	0%	Independent reading with prosodic manipulation	2 syllables with competing /l/ or /w/
10	0%	Independent reading with prosodic manipulation	Words in carrier phrases
11	0%	Independent reading with prosodic manipulation	Words in sentences
12	0%	Independent reading with prosodic manipulation	Sentences with multiple /r/ targets

Note. Parameters (represented in columns) change on a rotating basis between levels; the parameter that was changed in a given level is in bold.