

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/283036250>

How To Get Started With Ultrasound Technology for Treatment of Speech Sound Disorders

Article in *Perspectives on Speech Science and Orofacial Disorders* · October 2015

DOI: 10.1044/ssod25.2.66

CITATIONS

6

READS

1,410

3 authors, including:



[sue ann Lee](#)

Texas Tech University

36 PUBLICATIONS **277** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



the treatment efficacy of multiple opposition [View project](#)

How To Get Started With Ultrasound Technology for Treatment of Speech Sound Disorders

Sue Ann S. Lee

Texas Tech University Health Sciences Center
Lubbock, TX

Alan Wrench

Queen Margaret University
Edinburgh, UK

Sherry Sancibrian

Texas Tech University Health Sciences Center
Lubbock, TX

Disclosures: Financial: Sue Ann Lee and Sherry Sancibrian have no financial interests to disclose. Alan Wrench is associated with Articulate Instrument Ltd.

Nonfinancial: Sue Ann Lee presented part of this material at the 2014 ASHA Convention.

Alan Wrench presented part of this material at 2015 Ultrafest VI. Sherry Sancibrian has no nonfinancial interests to disclose.

Abstract

Although providing visual feedback with ultrasound is becoming more and more popular among speech-language pathologists, it is still not widely used, perhaps because many speech-language pathologists are unfamiliar with its application to treatment of speech sound disorders. This report introduces basic information on ultrasound for new users and summarizes existing studies conducted with ultrasound to provide clinical evidence for speech treatment. This report also presents a clinical case using ultrasound and discusses a treatment protocol for establishing postvocalic /r/ to illustrate an example of how ultrasound was used in a clinical setting.

Introduction

In treatment of speech sound disorders, patients may receive visual feedback for their speech gestures in various ways. For example, a mirror is traditionally used to allow patients to see their lip and tongue gestures. A variety of technologies such as acoustics and electropalatography also provide spectral or physiological feedback to help patients understand their tongue gestures and contact positions. Recently, ultrasound has become a new way to provide patients with real-time visual feedback of tongue position and configuration.

There are several advantages to use of ultrasound for treatment of speech sound disorders compared to other visual feedback resources. First, linear predictive coding (LPC) spectrum or spectrogram displays using acoustic software (e.g., Sona-Match by Kay Pentax) provide indirect feedback on tongue gestures. Patients must be equipped with knowledge of acoustics-articulation correlates to understand how the spectrum or spectrogram relates to their tongue gestures. Ultrasound, on the other hand, easily enables patients to identify their real-time tongue configuration and movement if they understand how the tongue is presented on either sagittal or coronal views of the ultrasound image. Neel (2010) presented the use of acoustic phonetics in clinical practice in *Perspectives on Speech Science and Orofacial Disorders*. It will be a useful resource if anyone is interested in learning how acoustics can be applied to clinical practice. Second, electropalatography, a tool showing location of the tongue's contact on the hard palate, provides physiological feedback

of tongue position. Electropalatography would be a useful clinical tool, but it may be less beneficial for young children because patients must wear an artificial plate made from a dental impression. The artificial plate placed against the palate may make young patients feel uncomfortable during speech. Gibbon and Lee (2007) provided information on the use of electropalatography as a clinical tool in *Perspectives on Speech Science and Orofacial Disorders*. Ultrasound is less invasive than electropalatography because the ultrasound probe is placed externally underneath the chin. However, gel added to the ultrasound probe may be uncomfortable for young children who have high tactile sensitivity. More specific information on how to use ultrasound will be addressed later.

Stone (2005) suggested that ultrasound has become more popular because it is non-invasive, low-cost, portable, and easy to operate. With all of the advantages, it is expected that more and more speech-language pathologists will use ultrasound technology for treatment of speech sound disorders in the future. However, only a limited number of speech-language pathologists have utilized ultrasound technology for speech treatment. Partly, this may be due to the fact that most existing literature focuses on disseminating research findings using ultrasound, rather than introducing the use of ultrasound technology for new users. Therefore, the main purpose of this current paper is to provide new users with basic knowledge of ultrasound technology, and to answer practical questions new users may raise when they start using ultrasound for their clinical practice. For experienced users to learn more comprehensive information regarding ultrasound, the published works by Hedrick, Hykes, and Starchman (1995) and Stone (2005) are recommended.

Ultrasound

Ultrasound refers to high frequency sound that is beyond the upper limits of the human hearing range. Humans hear sounds from 20 Hz to 20 KHz while diagnostic ultrasound operates at 1–20 MHz. Unlike human beings, animals such as bats and dolphins produce ultrasound and listen to returning ultrasound echoes to navigate or to identify objects.

Ultrasound has been used for several human applications in various fields (Society of Ultrasound in Medical Education, 2015). In 1914, the sinking of the Titanic led to the first ultrasound machine to detect icebergs. Later, ultrasound was employed to detect submarines in the ocean. In the 1940s, ultrasound was first used in the medical field by researchers to characterize tissues of internal organs. In 1963, the first commercially available two-dimensional ultrasound scanner was developed by Siemens.

In diagnostic medical ultrasound, reflected sound produces a dynamic image that reflects real time movements of the targeted anatomy. During 2D (B-mode) imaging, a transducer generates sound waves by applying carefully timed electrical pulses to a subset of the available piezoelectric crystals in the tip of the transducer. As these sound waves encounter parts of the anatomy, they are reflected back and received by the transducer. These echoes are then converted into a computer image allowing patients to view organ structure or movement in real time. Diagnostic ultrasound frequency ranges from 1–20 MHz, but 2–7 MHz ultrasound is most often used for images of lingual gestures.

An ultrasound machine is composed of the following parts: Transducer, Central processing unit (CPU), Transducer pulse controls, Display monitor, Keyboard, Storage device, and Printer (Freudenrich, 2001). The transducer (also called probe) and transducer pulse controls are essential parts of the ultrasound instrument. The transducer sends or receives sound waves. Transducer pulse controls change the amplitude, frequency, and duration of the sound wave emitted from the transducer. An expensive ultrasound machine contains all of these parts as a unit (see Figure 1). On the other hand, a relatively inexpensive ultrasound system is composed of probes attached to a personal computer with ultrasound software installed (see Figure 2). The more expensive ultrasound systems do not necessarily provide better quality images. They tend to have more advanced functions such as color Doppler and more than 2-dimensional imaging but

these advanced functions are not commonly used for speech therapy. Regardless of the expense of a particular system, factors such as transducer frequency, field of view, image frame-rate, and ease of recording the images with synchronous audio need to be considered.

Figure 1. Ultrasound Machine as a Unit.

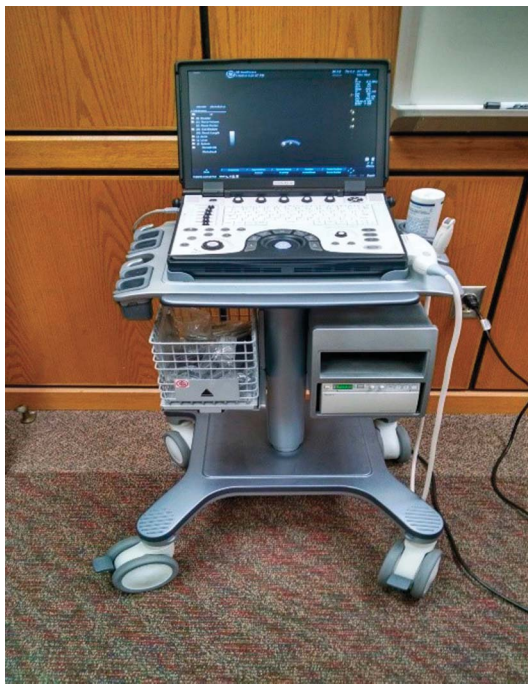


Figure 2. Ultrasound Probe Attached to a Computer.



When speech-language pathologists choose an ultrasound system for speech intervention, several aspects should be considered. First, the frequency ranges and shape of the probes, length of the probe handle, the radius of the probes and field of view (FOV) should be considered when probes are selected. Not all ultrasound models will have suitable probes. The frequency of emitted sound waves determines how deep the sound waves penetrate and the resolution of the image. Diagnostic ultrasound frequency ranges from 1 M Hz to 20 M Hz. Low frequencies can penetrate deeper areas of the body whereas high frequencies are used for detecting superficial structures. Frequencies from 2–7.5 MHz ultrasound have been used for ultrasound biofeedback treatment.

More specifically, a probe with frequencies from 2–4 M Hz may be required to image the whole tongue for adults whereas a probe with higher frequency (5–7 M Hz) may be adequate to image children's tongues. There are linear, curvilinear, convex, and micro-convex (<20mm radius) transducers. Because the probe is placed underneath the chin in ultrasound biofeedback treatment, convex or micro-convex probes with small radius are commonly used for midsagittal tongue imaging. The FOV refers to areas depicted by the ultrasound. The large field of view provides wider view of the tongue. Most ultrasound machines have probes available with 90° to 150° FOV. Probes with long handles may not be ideal as flexibility in the orientation of the probe is impeded by the handle meeting the chest.

A second important aspect is the size, weight, and portability of the ultrasound machine best suited for assessment and treatment. For clinical purposes, the ultrasound unit often needs to be moved to different clinic rooms or locations. Thus, small and lightweight would be ideal. A number of manufacturers produce ultrasound. Ultrasound systems made by Aloka, Mindray, GE, Sonosite, Terason, Ultrasonics, and Articulate Instruments have been used in speech lab and clinics. There are several portable ultrasound systems available on the market.

In order to help new users decide which ultrasound machine and probes to choose, Table 1 shows examples of ultrasound systems and types of probes used in several previous studies for either clinical or research purpose. There is no distinction among types of probes for either research or clinical purpose. If the same ultrasound was used by the same research group for multiple applications, only the equipment identified in the primary study for each research group was included in the table. It should be noted that this list does not contain all ultrasound machines used in all ultrasound research as well as clinical studies. Some of ultrasound systems (see *) are no longer available in the market.

Table 1. Examples of Ultrasound Systems and Probes Used in Previous Studies

Published study	Ultrasound	Probes	Types of study
Bernhardt et al. (2003), Bernhardt, Gick, Bacsfalvi, & Adler-Bock (2005)	Aloka SSD-900*	3.5 M Hz Convex Intercostal transducer	Clinical
Bernhardt et al. (2005)	Aloka Pro-Sound SSD 5000*	6 M Hz transducer	Clinical
Bernhardt et al. (2005)	Sonosite 180 Plus*	2-4 M Hz C15 Convex MCX transducer	Clinical
Bressmann et al. (2005); Bressmann, Flowers, Wong, & Irish (2010)	GE Logiq Alpha 100 MP*	6.5 M Hz E72 Micro Convex	Clinical & Research
Davison (2006). Klein et al. (2013)	Sonosite Titan*	5-8 M Hz C11 transducer	Research
Zharkova, Hewlett, & Hardcastle (2011)	Mindray DP6600*	65EC10EA	Research
Cleland, McCron, & Scobbie (2013)	Ultrasonix RP*	5-9 M Hz C9 Micro Convex	Research
Song, Demuth, Shattuck- Hufnagel, & Menard (2013)	Sonosite 180 Plus*	7-4 M Hz C11	Research
Preston et al. (2013) Byun et al. (2014)		Seemore PI 7.5M Hz Probe**	Clinical
Lin & Demuth (2015)	Terason t3000	3-8 M Hz 8MC3 Micro Convex transducer	Research
Cleland et al. (2015)	Ultrasonix SonixRP*	5-9 M Hz C9 Micro Convex	Clinical
Current study (2015)	GE Logiq e	8C-RS	Clinical

Note: *Discontinued models ** Only probe attached to a personal computer

Recently, handheld ultrasound scanners are starting to appear in the market (e.g., GE Vscan). Some ultrasound systems have displays that are too small to be useful for speech treatment, whereas some do not have appropriate probes for scanning lingual configurations. These systems are not capable of recording scanned images and motion figures. Thus, they may not be the best choice for biofeedback ultrasound treatment with the current level of functioning. Nevertheless this may be a type of device that has potential in the future if these issues can be resolved.

During treatment, patients are seated in front of the ultrasound unit (either personal computer with probe or ultrasound machine) so that they are looking directly at the monitor. Before scanning, gel is placed on the rim of the probe. As a contact medium between the skin and the head of the ultrasound transducer, the gel helps sound waves travel from the ultrasound head to skin without crossing through air. Because detecting tongue position and configuration is of interest in speech treatment, the probe is positioned in the medial position beneath the patient's chin. The surface of the probe must be in contact with the patient's skin in order to transmit ultrasound into the body. By rotating the probe 90 degrees, ultrasound provides two views. A midsagittal view shows the tongue configuration from front to back (see Figure 3). In Figure 3, the light arch-shape in the middle of the figure shows the surface of the tongue. The back of the tongue is to the left, and the front of the tongue is toward the right. The coronal view shows the

tongue configuration from right to left side of the front side of the tongue. The back side of the tongue is not shown with the coronal view (see Figure 4).

Figure 3. Sagittal View.

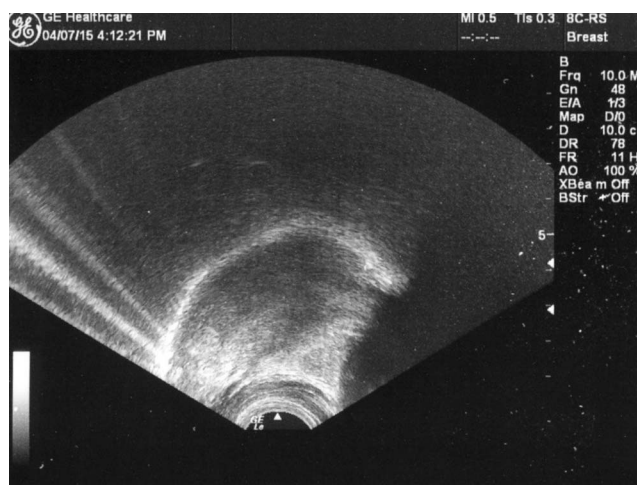


Figure 4. Coronal View.



Recording Image Sequences and Synchronizing With Audio

As well as using the ultrasound image for visual feedback, it is useful to be able to record and review ultrasound tongue movies both for diagnostic purposes and for tracking progress in therapy. Such recordings may result in the adjustment of original auditory-only diagnoses and help identify contexts that can be used to stimulate the target speech sound.

Nearly all ultrasound systems will allow image sequences to be saved and reviewed but many do not have the inbuilt facility to record the speech acoustics at the same time and without synchronized audio it is difficult to identify with certainty the region of the movie that corresponds to the speech sound of interest.

Ultrasound movies have been generated in a number of ways by researchers. A video recorder can be used to capture the picture output from the ultrasound machine's video output port while a microphone input to the video recorder captures the sound at the same time. Such a video recorder may be a stand-alone unit or may be run from a PC with a dedicated video card and software. Another approach that works for ultrasound systems that use a PC to provide the display is to run screen capture software such as CamStudio. The frame rate and synchronization accuracy of these two approaches is likely to be adequate for speech therapy review although not recommended for detailed study of speech dynamics where more precise alignment is required. A third approach adopted by a couple of research groups is to take specific ultrasound systems and modify the display software so that they record synchronous audio (Cleland, Scobbie, & Wrench, 2015; Hueber, 2013). Integration of the recording into the ultrasound software simplifies the process of recording and playing back movies as well as providing more accurate synchronization.

Safety

Although ultrasound is widely considered to be a safe technique (National Council on Radiation Protection, 2002) it is a form of energy which, above certain thresholds, can cause tissue heating and/or interaction effects with microscopic bubbles in the body (Health Canada, 2001); a cautious approach should therefore be taken in its use.

Dwell time is the amount of time a transducer remains in one place. Unlike typical diagnostic ultrasound exposure, which is nearly always very short in duration, an ultrasound therapy session may last 20 minutes or more. Also unlike typical ultrasound applications where the probe is moved around, when used for speech therapy, the probe remains more or less in the same position under the chin. These factors combine to result in a dwell time for speech therapy which is significantly higher than most other uses.

The second factor regulating ultrasound exposure is the power output of the ultrasound machine (sometimes referred to as "acoustic power"). Nearly all ultrasound machines have a controller that regulates the output power. Increasing power output increases the intensity of the beam and therefore the strength of echo return to the transducer; however, it also increases the patient's ultrasound dose. It is best practice, therefore, to operate on minimum power and maximum gain. Because turning up the gain control cannot compensate for insufficient power, a better solution is to use a lower frequency transducer if the image fades at depth.

To keep the ultrasound dosage as low as possible the following steps should be considered.

1. Use a probe with a frequency that is sufficiently low to image your client so that the power can be set to minimum with no loss of image.
2. Set the output power of the ultrasound machine to its minimum, turning up the gain if necessary to compensate.
3. Freeze the ultrasound display when not recording or actively using the visual feedback. When in freeze mode ultrasound is not generated.

It should also be noted that if an ultrasound probe is left running (i.e., not in freeze mode) while not in contact with tissue then the surface of the probe can heat up. Care should be taken to avoid this and to ensure that it is cool before applying it to the client.

A separate but no less important issue is care of the probe. If the probe is accidentally dropped or incurs any other trauma, it should be carefully inspected for cracks and if any are found, then the probe should not be used. In addition, care should be taken when cleaning the probe as some parts of the probe can deteriorate on contact with certain cleaning agents. Always refer to the manufacturer's instructions for advice on cleaning the transducer.

The Use of Ultrasound in Treatment of Speech Sound Disorders: Clinical Evidence

Since the 1980s, speech-language pathologists have reported using ultrasound in speech treatment with various types of patients in order to evaluate its effects. Early studies were conducted by Shawker and Sonies (1985) and Wein, Böckler, Klajman, and Obrebowski (1991). A significant body of early work was performed by Bernhardt and her colleagues. Bernhardt, Gick, Bacsfalvi, and Ashdown (2003) provided 14 weeks of treatment using two visual feedback technologies—electropalatography and ultrasound imaging—to four adolescents with moderate to severe hearing loss. Treatment targets included fricatives and liquids (/s, ʃ, r, l/) as well as the tense-lax high vowel contrast (/i, ɪ, u, ʊ/). As non-treatment targets, /t, d, ʈ/ and /e, ε, ə, ɜ:/ were selected. They found significant improvements for target sounds in production accuracy compared to non-treatment targets during post-treatment evaluation. The follow-up study (Bacsfalvi & Bernhardt, 2011) revealed that five out of seven participants generalized or maintained their level of performance, suggesting that speech treatment with visual feedback appeared to have lasting effects. The subsequent studies (Adler-Bock, Bernhardt, Gick, & Bacsfalvi, 2007; Bernhardt et al., 2008; Modha, Bernhardt, Church, & Bacsfalvi, 2008) confirmed the effectiveness of ultrasound in other types of patients in that use of ultrasound led to improvement of /r/ production in school-aged children, ages 7 to 15 years, with residual speech errors.

Recently, Preston, Brick, and Landi (2013) evaluated the efficacy of ultrasound biofeedback treatment for children with childhood apraxia of speech (CAS). Six children ages 9–15 years received 18 sessions of speech therapy. The target sounds varied depending on the children. They included initial consonantal /r/, initial /r/ or /l/ clusters, /ɜ:/ and other types of vocalic vowels (e.g., /ar/). The findings of the study revealed that the percentage of phoneme accuracy improved in 5 out of 6 participating children with CAS post-treatment. Accurate productions were maintained 2 months after treatment. In another study, Preston and his colleagues (Preston et al., 2014) used ultrasound for 7 children and 1 adult with residual speech sound errors in their motor-learning-based treatment program. The study found increases in production accuracy of target sounds in most participants. Generalization to clusters as well as to sentence-level production was observed. A 2-month follow-up showed evidence of retention. The authors argued that a motor-learning treatment program with ultrasound visual feedback can help individuals with residual speech sound errors.

Byun, Hitchcock, and Swartz (2014) also examined the effect of ultrasound biofeedback treatment for consonant /r/ and /ɜ:/ production in two studies. In the first study, 4 children, age range from 6 to 10, received speech treatment for 8 weeks. The authors failed to find a strong effect of ultrasound in improving /r/ production. Although 3 children correctly produced /r/ sounds with ultrasound visual feedback, there was little generalization effect when no visual feedback was available. In their second study, another 4 children, age range from 7 to 15, received ultrasound biofeedback treatment. Unlike the first study, the researchers found a large treatment effect across all participants. These gains were generalized to untreated words without visual feedback using ultrasound. The authors reported protocol differences between the two studies: in Study 1, all children were cued to match a bunched /r/ tongue configuration whereas in Study 2, children were cued to produce perceptually correct /r/, regardless of bunched or retroflex /r/ configuration. Hitchcock and Byun (2015) also reported additional evidence of the ultrasound effect in a case study and confirmed their previous findings.

Cleland et al. (2015) published a set of case studies on 7 children (aged 6–11) with a variety of persistent speech sound disorders. Standardized tests carried out on these children revealed 14 processes requiring attention but after eliminating structural processes and processes not related to changes in place of articulation, five error types were identified as potentially suitable for ultrasound biofeedback treatment: gliding, velar fronting, postalveolar fronting, backing, and idiosyncratic /r/ production. The process with the highest number of errors, excluding gliding,

was chosen as the initial target for therapy. Four children were treated for velar fronting; one for post-alveolar fronting; one for idiosyncratic backing (alveolars were produced as voiceless palatal lateral affricates); and one for idiosyncratic /r/ productions (/r/ was realized as a palatal lateral approximant). After intervention, all the children made significant progress on targeted segments, evidenced by both perceptual measures and changes in tongue-shape.

In summary, the previous studies consistently reported the usefulness of ultrasound technology for treatment of speech sound disorders. However, types of patients were limited to adolescents and adults with hearing loss, children and adults with apraxia of speech, or children with residual speech errors. Although Bernhardt et al. (2003) examined a greater variety of sounds such as fricatives /s, ʃ/ and liquid /r, l/ as well as high tense-lax vowels, and Cleland et al. (2015) treated a variety of disordered processes, the treated target sounds in the majority of studies have been limited to consonantal and vocalic /r/. Thus, more studies are warranted to evaluate the effectiveness of ultrasound in more diverse clinical populations as well as various target sounds.

Clinical Case Example

Although the previous studies demonstrated the positive effect of ultrasound use in speech therapy, there is still an ongoing need for examining the effect of ultrasound in a variety of patients. In this paper, we provide a case example using ultrasound. Because it is a retrospective description of a real case in a working clinic at Texas Tech University Health Sciences Center, this clinical case is not as well-controlled as other clinical research studies were. However, it is one example showing how a basic clinical case study using ultrasound can be conducted in a clinical setting. For more rigorous clinical research using ultrasound, see other published research papers addressed above.

This case example showed the effect of ultrasound in improving postvocalic /r/ production in an adolescent with residual errors. Approximately 6% of school-age English-speaking children demonstrate speech production errors (Williams, 2003). It is well-known that the English /r/ is a challenging sound for children to acquire. This difficulty is in part due to the complexity of the articulatory gesture to produce the /r/ sound. In English, the articulation of /r/ generally falls into two categories (Shriberg & Kent, 2002) in terms of the anterior portion of lingual constriction. Some speakers may raise the tongue tip and curl it back to approximate a point behind the alveolar ridge (i.e., retroflex /r/). Others may lower the tongue tip, but raise the tongue blade toward the palatal area (i.e., bunched /r/). It is possible to observe variations of tongue placement or gestures across native English speakers (Delattre & Freeman, 1968). Although the tongue configurations may differ across individuals, these variants are perceptually the same and the first three formant frequencies are acoustically similar (Zhou et al., 2008). Ladefoged (2003) reported that the retroflex /r/ is the most common followed by the bunched /r/. However, others (Boyce et al., 2009; Mielke, Baker, & Archangeli, 2010) found that more native speakers of American English produced /r/ with the bunched shape than the retroflex /r/. All claim that speakers also produce /r/ with varying tongue shapes, including possible intermediate positions. When /r/ is produced before vowels, it is called prevocalic or consonantal /r/ (e.g., *run* and *rock*) whereas /r/ following a vowel creates what is known as a rhotic diphthong (e.g., *ear*, *air*, *door*, *poor*, and *car*). In addition to the anterior part of lingual constriction, two additional gestures are involved in /r/ production; lip rounding and constricted tongue root that comes close to the pharynx (Adler-Bock et al., 2007). Klein, Byun, Davison, and Grigos (2013) emphasized the importance of constriction in the pharyngeal regions to produce perceptually correct /r/.

Patient

The patient in this case example was a 13-year-old native monolingual English-speaking male who had been diagnosed with mixed expressive-receptive language disorder, developmental articulation disorder, developmental reading disorder, and auditory processing disorder. Although he had been receiving speech-language treatment at a university clinic since August 2009, his misarticulations continued to persist. The error productions included most notably vocalic /r/,

fricative /s/, and affricate /tʃ/. All vocalic /r/ sounds were produced without r-color (rhotacization). His /s/ sound was lateralized. The voiceless affricate consonant /tʃ/ was produced as /ʃ/. He also had difficulty producing and sequencing multisyllabic words. An examination of the participant's oral mechanism indicated that all structures and functions were adequate to support speech. His hearing sensitivity was within normal limits bilaterally. Overall intelligibility to familiar listeners in an unknown context was approximately 80%. Because his intelligibility was judged as only mildly impaired, his regular therapy focused on his language and literacy skills. No other speech goals were addressed while the present treatment was being provided.

Target Sounds

The major error in his speech production was vocalic and postvocalic /r/. Thus, all forms of vocalic and postvocalic /r/ such as /ɜr ɪr ɛr ɔr ʊr/ were targeted. In addition, /tʃ/ was selected as a non-treatment sound. If the treated sounds except for /tʃ/ were improved, it indicates that the improvement of vocalic and postvocalic /r/ resulted from ultrasound treatment.

Treatment Procedures

Treatment was provided by a Texas licensed speech pathology intern and took place in a quiet room at a university clinic. A licensed and certified SLP supervised her and offered weekly consultation for improving treatment fidelity. Prior to the first session, baseline data were collected using a picture-naming task. Five probes of a one-syllable word for each postvocalic /r/ and /tʃ/ sound were adopted, resulting in a total of 35 probes. After the baseline session, a 30-min individual intervention session was provided once a week for 12 weeks. During treatment, various vocalic and postvocalic /r/ sounds in one-, two-, and three-syllable words were used. The specific treatment protocol will be addressed below. Once the 12 weekly sessions were completed, a post-treatment session was conducted with the same probes used for baseline.

GE Logic-e Ultrasound with an 8C-RS transducer was used to provide visual feedback during treatment sessions. This is a high-quality ultrasound machine that provides a clear image of the tongue position and configuration. The patient was seated in front of the ultrasound unit so that he looked directly at the monitor to see his tongue configuration. Then, the probe was positioned in the medial position beneath the chin. Typically, a midsagittal view was chosen where the tongue configuration from front to back was shown on the monitor. When the probe was angled in an anterior position, the tongue blade and tip were clearly visible. On the other hand, the tongue dorsum was visible when the probe was angled in a posterior position.

Treatment Protocol

Ultrasound was used through all 12 weeks of treatment with varying degrees of use for visual biofeedback. The first five sessions, ultrasound was used for the entire treatment session. Once the patient was able to make the retroflex /r/ lingual gesture and produce an /r/-color sound, various table-top activities were introduced to elicit word productions. Ultrasound was not used during table-top activities.

During the first treatment session, the patient was asked to produce /ɜr/ and watch his real-time tongue movement on the monitor. The patient was not able to make a correct gesture. The ultrasound imaging revealed that his tongue was consistently flat and stayed in mid-central position of the oral cavity without any additional tongue tip or dorsum configuration. Perceptually, his production sounded like /ʌ/. Next, the clinician demonstrated and produced /ʌ/ and /ɜr/ (see Figures 5 & 6). After the clinician's and patient's static as well as dynamic movement images were saved into the ultrasound machine, the images played over and over again until the patient understood where the tongue tip, tongue blade, and tongue dorsum were shown on the monitor. Also, the clinician's and patient's images were compared to each other so that the patient understood the differences between the two speakers.

Figure 5. Clinician's Model for /ʌ/.



Figure 6. Clinician's Model for /ɜː/.



In order to identify which tongue gesture to teach, the clinician asked the patient to make attempts to produce two different gestures of /r/ (i.e., retroflex vs. bunched). It seemed that he felt more comfortable with the articulation of retroflex /r/. Thus, the articulation of retroflex /r/ was selected for the target articulation.

During the second treatment session, the procedure from the first session was repeated. In addition, a small amount of chocolate paste was applied to the post alveolar-palatal position. The patient was asked to approximate the chocolate paste with the tongue tip raised and curled back. His gesture was monitored by ultrasound. The imaging indicated that he was able to raise his tongue tip and curl back to approximate the chocolate paste. It should be noted that the patient was not encouraged to sweep the tongue along the hard palate to lick the paste or touch the hard palate in order to prevent any incorrect gesture. During this training, the patient was not asked to produce any phonation. After several repetitions, the patient was able to curl his tongue tip back behind the alveolar ridge without phonation.

During the third session, the same procedure as in the first and second session was repeated. A small amount of chocolate paste was still applied. Although he was able to curve his tongue back, he was not able to maintain the curled tongue position when he attempted to produce the /ɜ:/ sound synchronously.

During the fourth session, ultrasound imaging was used to view the patient's tongue gestures while he practiced moving from a relaxed position to a curled tongue position. We continued to use the chocolate paste. He was again asked to produce the /ɜ:/ sound. After several attempts, the patient started to synchronize the sound and gesture. However, his productions lacked the rhotacization (r-coloring) required for /ɜ:/.

During the fifth session, the patient was asked to produce /ʌ/ and /ɜ:/ alternatively so that he could see how his tongue changed from one position to the other on the ultrasound monitor. No more chocolate paste was applied. We used the word "her" as a target.

During the sixth session, the patient was able to produce /ʌ/ and /ɜ:/ contrastively. We used phrases like "her shoes, her book, and her desk" in order to generalize his production beyond the single word level without ultrasound. Then, we introduced other rhotic diphthongs such as /ɪr/ and /ar/ using ultrasound.

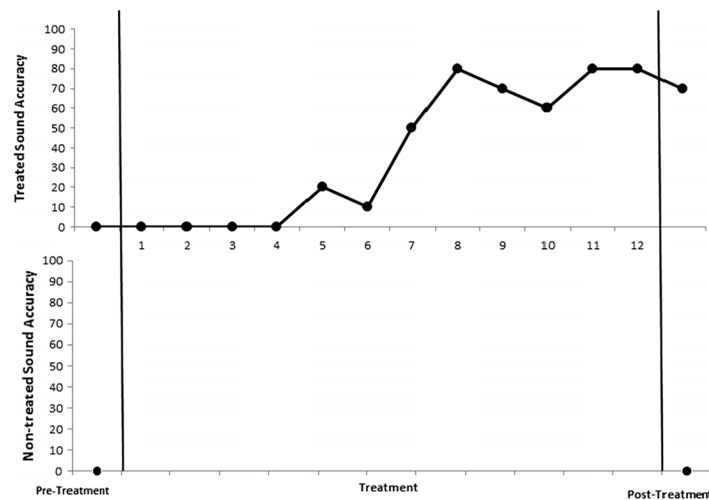
During the seventh through ninth sessions, the patient continued to work on /ɜ:/, /ɪr/ and /ar/. His gestures were monitored by ultrasound at the beginning of each session. Then, we incorporated various tabletop activities using picture cards (e.g., go fish game, bingo game) without ultrasound to practice these target words for generalization.

During the tenth through twelfth sessions, additional rhotic diphthongs /ɛr, ʊr, ɔr/ were treated. Ultrasound continued to be used to monitor the tongue gestures. Various tabletop activities without ultrasound were also incorporated.

Treatment Progress

Figure 7 shows the percent accuracy of target vocalic /r/ as well as the non-treatment target /ʃ/. During the baseline, the production accuracy of both postvocalic /r/ and /ʃ/ sounds was zero. All vocalic and postvocalic /r/ productions lacked r-coloring. The /ʃ/ sound was replaced by /ʒ/. The baseline accuracy of zero percent continued for /ɜ:/ production during the first four sessions. However, production accuracy slightly increased to 20% in the fifth session. In the sixth session, production accuracy decreased to 10% in part due to the fact that the /ɜ:/ sound was targeted in phrases. However, production accuracy increased to 50% in the seventh session with new vocalic /r/ sounds. The remaining sessions showed consistently increasing accuracy even after the other new rhotic diphthongs were added.

Figure 7. Production Accuracy of Vocalic /r/ Across Sessions.



A post-treatment test was given using the same probes after 2 weeks. The post-test revealed that 70% of vocalic and five postvocalic /r/ forms were correctly produced. The non-treatment target /ɹ/ was also evaluated. The /ɹ/ sound continued to be produced as /ʃ/ in a picture-naming task.

Discussion of Clinical Case Example

Most previous studies reported the positive effects of ultrasound in treatment of speech sound disorders (Adler-Bock et al, 2007; Bernhardt et al., 2003, 2008; Bacsfalvi & Bernhardt, 2011; Byun et al., 2014; Modha et al., 2008; Preston et al., 2013, 2014). Consistent with previous studies, the present case also revealed an improvement in vocalic /r/ production in an adolescent patient. Although the patient had a long history of incorrect production of vocalic /r/ sounds, the patient was able to learn the correct tongue gestures and produce the target sounds with 70% accuracy within a 12-week period of intervention.

It seems that ultrasound is particularly useful at the beginning of treatment as a tool to teach patients what their actual tongue gesture looks like. The patient in this study was engaged and motivated by viewing his own tongue shape and gestures. However, it is also important to note that it is necessary to incorporate regular tabletop activities in therapy while visual feedback using ultrasound is being provided because the patient may lose interest if therapy only relies on ultrasound feedback.

There are several limitations of our case example due to lack of experimental control. Only one patient participated in this case study. No replication across participants or behaviors was made. It should be also noted that besides the anterior gesture, posterior constriction should be stressed during treatment as Klein et al. (2013) suggested. Furthermore, it is recommended that the coronal view in addition to the midsagittal view of ultrasound display be also used during treatment (Adler-Bock et al., 2007). Finally, although the present case provided positive evidence of improvement during the therapy period, it is not certain whether or not the patient continued to maintain the same level of performance after dismissal.

Conclusion

The present paper was designed to introduce how ultrasound can be used in treatment of speech sound disorders and provided a summary of clinical evidence using ultrasound for new

users. Previous research papers using ultrasound focused on disseminating research findings. There are few resources available specifically for new users who want to use ultrasound technology for speech treatment. It is our hope that this material provides a basic guideline for getting started with ultrasound and encourages speech-language pathologists to take advantage of this emerging technology.

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., & Bernhardt, B. M. (2011). Long-term outcomes of speech therapy for seven adolescents with visual feedback technologies: Ultrasound and electrophalatography. *Clinical Linguistics & Phonetics*, 25, 1034–1043.
- Bernhardt, B. M., Gick, B., Bacsfalvi, P., & Adler-Bock, M. (2005). Ultrasound in speech therapy with adolescents and adults. *Clinical Linguistics & Phonetics*, 19, 605–617.
- Bernhardt, M. B., Bacsfalvi, P., Adler-Bock, M., Shimizu, R., Cheney, A., Giesbrecht, N., . . . Radanov, B. (2008). Ultrasound as visual feedback in speech habilitation: Exploring consultative use in rural British Columbia, Canada. *Clinical Linguistics & Phonetics*, 22(2), 149–162.
- 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.
- Boyce, S. E., Tiede, M. K., Groves-Wright, K., Espy-Wilson, C., Holland, C., & Choe, K. (2009). Dialect distribution of North American English /r/. *The Journal of the Acoustical Society of America*, 125, 2575.
- Bressmann, T., Thind, P., Uy, C., Bollig, C., Gilbert, R. W., & Irish, J. C. (2005). Quantitative three-dimensional ultrasound analysis of tongue protrusion, grooving and symmetry: Data from 12 normal speakers and a partial glossectomy. *Clinical Linguistics & Phonetics*, 19, 573–588.
- Bressmann, T., Flowers, H., Wong, W., & Irish, J. C. (2010). Coronal view ultrasound imaging of movement in different segments of the tongue during paced recital: Findings from four normal speakers and a speaker with partial glossectomy. *Clinical Linguistics & Phonetics*, 24, 589–601.
- Byun, T. M., Hitchcock, E. R., & Swartz, M. T. (2014). Retroflex versus bunched in treatment for rhotic misarticulation: Evidence from ultrasound biofeedback intervention. *Journal of Speech, Language, and Hearing Research*, 57, 2116–2130.
- Cleland, J., McCron, C., & Scobbie, J. M. (2013). Tongue reading: Comparing the interpretation of visual information from inside the mouth, from electropalatographic and ultrasound displays of speech sounds. *Clinical Linguistics & Phonetics*, 27, 299–311.
- Cleland, J., Scobbie, J. M., & Wrench, A. A. (2015). Using ultrasound visual biofeedback to treat persistent primary speech sound disorders. *Clinical Linguistics & Phonetics* [Advance Online Publication]. doi:10.3109/02699206.2015.1016188
- Davison, L. (2006). Comparing tongue shapes from ultrasound imaging using smoothing spline analysis of variance. *Journal of Acoustical Society of America*, 120, 407–415.
- Delattre, P., & Freeman, D. C. (1968). A dialect study of American r's by x-ray motion picture. *Linguistics*, 6, 29–68.
- Freudenrich, C. (2001). How ultrasound works. Retrieved from <http://science.howstuffworks.com/ultrasound.htm>
- Gibbon, F. E., & Lee, A. (2007). Electropalatography as a research and clinical tool. *Perspectives on Speech Sciences and Orofacial Disorders*, 17, 7–13.
- Health Canada. (2001). *Guidelines for the safe use of diagnostic ultrasound*. Retrieved from <http://www.hc-sc.gc.ca/ewh-semt/pubs/radiation/01hecs-secs255/index-eng.php>
- Hedrick, W. R., Hykes, D. L., & Starchman, D. E. (1995). *Ultrasound physics and instrumentation* (3rd ed.). St Louis, MO: Mosby Inc.
- Hitchcock, E. R., & Byun, T. M. (2015). Enhancing generalization in biofeedback intervention using the challenging point framework: A case study. *Clinical Linguistics & Phonetics*, 29, 59–75.

- Hueber, T. (2013). Ultraspeech-player: Intuitive visualization of ultrasound articulatory data for speech therapy and pronunciation training. *Proceedings of Interspeech, 2013*, 752–753.
- Klein, H. B., Byun, T. M., Davison, L., & Grigos, M. I. (2013). A multidimensional investigation of children's /r/ production: Perceptual, ultrasound, and acoustic measures. *American Journal of Speech-Language Pathology, 22*, 540–553.
- Ladefoged, P. (2003). *A course in phonetics* (5th ed.). Beverly, MA: Wadsworth Publishing.
- Lin, S., & Demuth, K. (2015). Children's acquisition of English onset and coda /l/: Articulatory evidence. *Journal of Speech-Language-Hearing Research, 58*, 13–27.
- Mielke, J., Baker, A., & Archangeli, D. (2010). Variability and homogeneity in American English /r/ allophony and /s/ retraction. *Laboratory Phonology, 10*, 699–730.
- Modha, G., Bernhardt, B. M., Church, R., & Bacsfalvi, P. (2008). Case study using ultrasound to treat /r/. *International Journal of Language and Communication Disorders, 43*, 323–329.
- National Council on Radiation Protection. (2002). *Exposure criteria for medical diagnostic ultrasound: II. Criteria based on all known mechanisms*, NCRP Report No. 140. Bethesda, MD: Author.
- Neel, A. T. (2010). Using acoustic phonetics in clinical practice. *Perspectives on Speech Sciences and Orofacial Disorders, 20*, 14–24.
- Preston, J. L., Brick, N., & Landi, N. (2013). Ultrasound biofeedback treatment for persisting childhood apraxia of speech. *American Journal of Speech-Language Pathology, 22*, 627–643.
- Preston, J. L., McCabe, P., Rivera-Campos, A., Whittle, J. L., Landry, E., & Mass, E. (2014). Ultrasound visual feedback treatment and practice variability for residual speech sound errors. *Journal of Speech, Language, Hearing Research, 57*, 2102–2115.
- Shriberg, L. D., & Kent, R. D. (2002). *Clinical phonetics* (3rd ed.). Boston, MA: Allyn & Bacon.
- Shawker, T. H., & Sonies, B. C. (1985). Ultrasound biofeedback for speech training. Instrumentation and preliminary results. *Investigative Radiology, 20*, 90–93.
- Song, J., Demuth, K., Shattuck-Hufnagel, S., & Menard, L. (2013). The effects of coarticulation and morphological complexity on the production of English coda clusters: Acoustic and articulatory evidence from 2-year-olds and adults using ultrasound. *Journal of Phonetics, 41*, 281–295.
- Society of Ultrasound in Medical Education (2015). Learning modules. Retrieved from <http://www.susme.org/learning-modules/learning-modules/>
- Stone, M. (2005). A guide to analyzing tongue motion from ultrasound image. *Clinical Linguistics & Phonetics, 19*, 455–501.
- Williams, L. (2003). *Speech disorders: Resource guide for preschool children*. Clifton Park, NY: Delmar, Cengage Learning.
- Wein, B., Böckler, R., Klajman, S., & Obrebowski, A. (1991). Ultrasonography of the tongue in the rehabilitation of speech articulation disorders. *Otolaryngologia Polska, 45*, 133–140.
- Wrench, A. (2013). *Ultrasound speech analysis: State of the art* [PowerPoint Slides]. Retrieved online from http://www.qmu.ac.uk/casl/conf/ultrafest_2013/programme.htm
- 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*, 4466–4481.
- Zharkova, N., Hewlett, N., & Hardcastle, W. J. (2011). Coarticulation as an indicator of speech motor control development in children: An ultrasound study. *Motor Control, 15*, 118–140.

History:

Received April 16, 2015

Revised August 01, 2015

Accepted August 06, 2015

doi:10.1044/ssod25.2.66