# The use of ultrasound biofeedback for improving English /r/

**Michelle Cavin**
University of Victoria
*mccavin3@gmail.com*

Research suggests that the use of a two-dimensional dynamic ultrasound machine is a viable option for improving speech sound production accuracy. This non-invasive biofeedback technology allows for a person to see the movement of their tongue-shape features, so they can then modify their own articulation to match a correct model. The purpose of this research is to evaluate the efficacy of ultrasound as a tool in speech therapy, more specifically for the remediation of the North American English /r/. A case study was conducted with a single 22-year-old male participant, consisting of nine recording sessions focusing on /r/ pronunciation. A comparison between the recordings from the first and last sessions showed a lowering of the third formant (F3) by 347.4 Hz, while the difference between F3 and F2 was 166.2 Hz lower post sessions. In addition, results from a Likert scale questionnaire given to three native speakers of English either agreed or strongly agreed to clearly hearing the /r/ sound 33.4% of the time in recordings from the first session, compared to 68.8% in recordings from the last session. Although the results of this study did not compare the use of ultrasound in therapy to traditional therapy, its findings add to the evidence that the use of ultrasound biofeedback can facilitate production accuracy, across a broader population than has previously been tested.
*Keywords: ultrasound; biofeedback technology; English /r/; lower third formant*

## 1 Introduction

Communication is, and always has been, vital in our day-to-day life. Not only that, but it is a basic right of all persons to communicate the conditions of their own existence (National Joint Committee, 1992). We use it to convey information to one another through both linguistic and non-linguistic forms, as well verbal and non-verbal. Although all forms can be effective, speech is the most common form of communication humans use. We are born with a capacity for language, as even babies who are born deaf babble as newborns (Berk, 2012). This necessity for communication, and more importantly speech in our very social society, affects the way children living with hearing and language impairments assimilate into classrooms. This also applies to adults who have suffered from health problems affecting their speech. The growing role of Speech Language Pathologists (SLPs) and the use of technology allow them to reach their full potential as members of society.

An example of this technology involves incorporating biofeedback treatment into speech therapy, as it is "a means of supplying an individual with information that is not normally available at a conscious level regarding the consequences of a behavior" (Shuster & Ruscello, 1995, p. 37). Research has proven technology that involves electromagnetic articulography, electropalatography, spectrograms, and ultrasound imaging to be effective, especially with persistent speech errors. These are errors that remain present from eight to nine years of age (Byun & Hitchcock, 2012). With that said, this technology is still recently being incorporated into clinical Speech-Language Pathology in British Columbia, as previous studies in this field are limited in number and scope. However, as Preston et al. (2013) state, "[T]hese alternative approaches might be viewed as clinically and economically viable if they can result in rapid and sustained gains in speech production" (p. 628).

The focus of the current paper and review of past studies will be on the use of the ultrasound imaging technology, which consists of a transducer and a monitor that displays a dynamic image of the tongue. The transducer is held beneath the speaker's chin, where ultrasonic waves meet with air in the oral cavity above the tongue and reflected back to the probe (Bernhardt et al., 2008, p. 150). According to Gick (2002), the ultrasound came into clinical use in the 1960s-70s, and since then it has been used in studies of speech production. Since then, ultrasound technology has advanced much like everything else and they have gone from being expensive and unwieldy to more affordable and even portable. This has opened many doors for the application of this useful tool, as it is more financially feasible over other forms of biofeedback and can ultimately be "implement[ed] on a larger scale clinical basis" (Preston, et at. 2013, p. 628).

In a study done by Adler-Bock et al. (2007), an ultrasound machine was used in the remediation of North American English /r/ in two hearing adolescents (12 and 14 years of age) with persistent /r/ difficulties. Prior to this study, no research was available on the use of ultrasound remediation of /r/ in this type of clients. Both children received 13 treatment sessions from an SLP that began with awareness and moved through a series of increasingly difficult activities for production of /r/. The participants had previously received traditional /r/ therapy with little success, and the results were measured through formant analysis and listener judgments rated by SLPs who were unfamiliar with the participants. They found an expected lowering of the third formant, and the post treatment ultrasound images of /r/ tokens "showed tongue shapes to be more similar to those of typical adults than had been observed before treatment" (Adler-Bock et al., 2007, p. 128). They concluded that alternative treatment methods like ultrasound imaging technology has "potential utility for remediation of /r/ in speakers with residual /r/ impairment" (Adler-Bock et al., 2007, p. 128).

In another recent study, Preston et al. (2013) included the ultrasound machine in treatment for children with "persistent speech errors associated with childhood apraxia of speech (CAS)" (p. 627). The researchers viewed the feedback of motor performance as a crucial part of the speech production system, which includes auditory and somatosensory information. When the speaker experiences this feedback, they can then make adjustments when errors arise. Individuals with CAS can experience a disruption in this process, which is why the use of biofeedback "may be useful for teaching children to recognize errors and adjust their productions" (Preston et al., 2013, p. 628). Six children participated in this study. The treatment focused on tongue movement sequences where researchers used both verbal and visual cues to elicit correct productions. All participants reached the pre-established criterion, which was 80% accuracy of production for two consecutive sessions of two treated sound sequences, in an average time span of five sessions. In a two-month follow-up, the researchers discovered that the participants maintained the progress they had made during treatment.

Although this study focused particularly on children with CAS, anyone with a speech sound disorder can benefit from biofeedback as a tool for providing visual information about nature of movement and target movement. As Preston et al. (2013) state quite eloquently, "[B]y teaching children about articulatory targets using visual feedback of tongue movements, and by sequencing these movements in various words/phrases/prosodic contexts, the relationship between the speech motor plan and the actual movements may be strengthened" (p. 628).

The goal of the current paper is to support and add to the limited literature available that suggests the use of biofeedback, in this case a two-dimensional dynamic ultrasound machine, as an effect tool for speech therapy. The current study was conducted by an undergraduate linguistics student interested in learning more about the use of ultrasound equipment in the field of Speech-Language Pathology. The purpose was to investigate how a person with a persistent /r/ difficulty can make changes to their articulation over a short time span when that person can see the shape and location of their tongue in real time. This was achieved through comparison of the student's and participant's articulation, and discussions about the environment where we see /r/ in the English language.

## 2 Methodology

### 2.1 Participant

An individual case study design was followed for the current study. The participant was a 22-year old Canadian English-speaking adult male, Jason. He grew up in a middle- to upper-class family in Smithers, British Columbia, where he was the youngest of three children. He never received speech therapy of any kind, and /r/ was the only speech sound that he had not acquired naturally. He self-reported having a hearing test at age eight, where he was told he had normal hearing. There were no other developmental concerns.

## 2.2 Procedure & Stimuli

The study comprised nine sessions, each one hour long. The first session consisted of an introduction to the ultrasound machine (GE Logiq-e with a 7X 8C-RS probe), as well as a recording of 16 target word tokens, found in (1), taken from Adler-Bock et al. (2007), as well as a discussion about the ultrasound machine. These words included /r/ in all the different environments found in English. These environments are word-initial, word-medial, word-final and consonant-cluster, where vowel and consonant contexts play a large role in the articulation of the following /r/ phoneme.

| (1) | *Word-initial* | *Word-medial* | *Word-final* | *Consonant -cluster* |
|---|---|---|---|---|
| | read | heary | ear | pray |
| | rid | hairy | air | bray |
| | ray | hurry | her | tray |
| | red | story | pour | dray |

The subsequent seven training sessions revolved around the discussion of /r/, as it is "one of the most complex phones in the English language in its articulatory and acoustic characteristics" (Adler Bock et al., 2007, p. 128). After watching the ultrasound video of the initial recording, the closest tongue configuration between tip-up retroflexed or tip-down bunched (Adler-Bock et al., 2007) for Jason's /r/ was judged to be tip-down bunched, therefore sessions revolved around this tongue movement. The visual screen of the ultrasound machine was positioned between the researcher and the participant, and sticky notes were placed on the screen to mark the tongue back, tongue body and tongue tip. This served as a reference point for Jason to ensure he was moving his tongue back and upwards towards the palate. The researcher's /r/ tip-down bunched /r/ articulation served as a model as comparisons were made.

An outline of the sessions is described in (2), as well as comments on his progress. Stimuli from the first seven sessions can be found in Appendix A, and sentences from session eight in Appendix B. The choice of stimuli across sessions was created with a difficulty hierarchy in mind, an idea taken from Preston et al. (2013). The 16 token words were not included in sessions, but instead words with similar environments. This way the token words were not practiced during sessions, and improvement was not bias to specific words.

(2)

| Session | Focus | Comments |
|---|---|---|
| 1 | Recording session of stimuli | Overall observation is there is not enough constriction taking place, replacement sound hard to classify. |
| 2 | Awareness of /r/ | Jaw is shut and needs reminders to keep slightly open. Raising tongue to palate is unnatural movement for him. /r/ sound achieved at the end. |
| 3 | Transitions involving /r/ | Jaw remains issue, pencil used as bite block, tongue needs to pull back and raise more. |
| 4 | Constrictions at the palate and pharynx | Not enough constriction being made with tongue, too much air passing through. Jason self-reports that constriction in pharynx is difficult and unnatural. |
| 5 | Transition of /r/ into vowels | Rounding is already taking place naturally, this transition is not difficult. Overall /r/ articulation becoming more natural. Markers on screen helping. |
| 6 | Transition of /r/ out of vowels | Difficult because of jaw movement, as well as overall tongue movement. Some vowels more difficult than others. Will review again next session. |
| 7 | Different vocalizations of /r / involving vowels | Reminder to articulate slower, as this is when best /r/ is achieved. Sound quality is largely improved. |
| 8 | Sentences with /r/ | Only three words in twenty sentences are found to be difficult. Articulation is still conscious effort. |
| 9 | Recording session of stimuli | Jason has made rapid gains in /r/ production accuracy. |

## 3 Results

The ultrasound image and audio recordings were synced using BlackMagic Design. The synchronization was needed to have Jason's ultrasound images match auditory information for later analysis. Results consisted of *qualitative* analysis of tongue contour (articulation) and *quantitative* analysis of formant values (acoustics) of /r/, pre- and post-training. In addition, auditory analysis was conducted in the form of a perceptual task administered to three native-English speakers, to determine to what extent observed improvements in /r/ were audible.

### 3.1 Articulatory Analysis

Ultrasound images of a frozen /r/ articulation are shown in Figure 1 and Figure 2; taken in Sessions 1 and 9 respectively. This first session image in Figure 1 shows Jason's tongue to be fairly relaxed, with no retraction of the tongue and constriction against the palate. A frozen ultrasound image taken from the last session image in Figure 2 shows more constriction between tongue and palate taking place, a more accurate tip-down bunched tongue configuration. Both stimuli are taken from the token "air", where the /r/ is in word-final position, and follows a vowel. This token was chosen as it allowed for a clear view of the /r/ articulation on its own after the vowel.

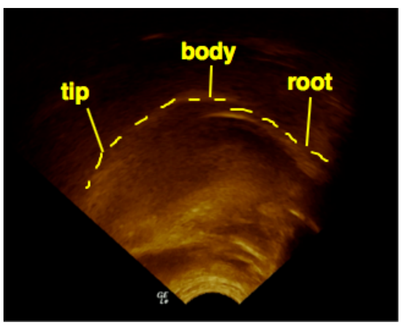


*Figure 1: First session /r/ articulation*

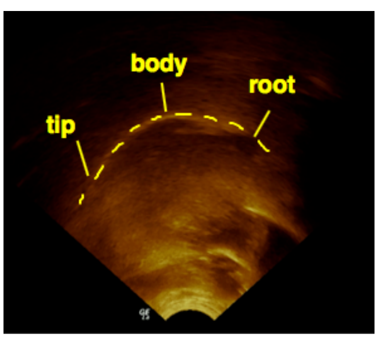


*Figure 2: Last session /r/ articulation*

### 3.2 Acoustic Analysis

To quantify the difference in articulation between the first and last sessions, the second, and third formants of /r/ were extracted from target words using a Praat script, and the averages were calculated across the 16 tokens. F1 and F2 are plotted using Excel. Values for F3 are provided in (3), as well as the difference between F3 and F2.

| (3) | *F3* | *F3-F2* |
|---|---|---|
| Pre | 2666.7 Hz | 1028 Hz |
| Post | 2318.3 Hz | 861.8 Hz |

As seen above, the frequency of F3 dropped between session one and session nine, consistent with improvement of /r/ production, as well as a smaller gap between F2 and F3 (Modha et al., 2008).

### 3.3 Auditory analysis

The auditory analysis consisted of a listening task that reflected the accuracy of the /r/ sound in 16 tokens between the first and last session. The tokens were presented in a randomized order, and participants were given a Likert scale questionnaire. Three native-English speakers completed this test, and their answers were consistent with each other. The results are presented below in (4), with Recording 1 taken from the first session, and Recording 2 from the last session.

| (4) | *I can clearly hear the /r/ sound in the word* | *Recording 1* | *Recording 2* |
|---|---|---|---|
| | Strongly Disagree | 10.4 % | 0% |
| | Disagree | 25% | 8.3% |
| | Neutral | 31.3% | 23% |
| | Agree | 29.2% | 50% |
| | Strongly Agree | 4.2% | 18.8% |

The word tokens from Recording 2 were consistently ranked as more accurate to the listeners, which also supported the idea that Jason's /r/ articulation previous to the study was perceived with low accuracy ratings. Although there are many different variants of /r/, Gick et al. (2003) state that "what differentiates the variants of /r/ is which part of the front of the tongue is used to execute the palatal constriction"(p. 76). With that said, they all must have some constriction taking place in that area, along with in the labial and pharyngeal regions. This means that although there can be acoustic differences in articulation, Jason's /r/ was not perceived as well articulated by other listeners.

## 4 Discussion

This study builds on previous work as it brings forth additional evidence on the benefit of visual feedback in clinical settings: qualitative and quantitative evidence indicated improvement of /r/ across the study. There was no period of traditional therapy with an SLP without the use of ultrasound involved in this investigation; therefore, it is not known whether improvements could have been made without the use of biofeedback. However, the participan
t made very rapid gains in /r/ production accuracy in a short time frame, after 22 years of misarticulation. The ultrasound proved to be especially useful when the target sound was being uttered. This study supports the idea that ultrasound is effective as a visual feedback tool in the remediation of /r/, and potentially other phones as well.

## Acknowledgements

I would like to thank Dr. Sonya Bird for her support and guidance, Dr. John Esling for his time and discussion of the first recordings, Chris Coey for all his technical assistance, and, of course, Jason for his time, effort, and enthusiasm.

## References

Adler-Bock, M., Bernhardt, B., 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*(2), 128–139.

Berk, L. E. (2012). *Infants and children: Prenatal through middle childhood.* Boston: Pearson Education.

Bernhardt, B., Bacsfalvi, P., Adler-Bock, M., Shimizu, R., Cheney, A.Giesbrecht, N., O'Connell, M., Sirianni, J. & 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.

Boersma, P. & Weenink, D. (2009). Praat: Doing phonetics by computer (Versions 4.5) [Computer program] http://www.praat.org/.

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.

Gick, B., Iskarous, K., Whalen, D. H. & Goldstein, L. M. (2003). Constraints on variation in the production of English /r/. In S. Palethorpe & M. Tabain (Eds.), *Proceedings of the 6th International Seminar on Speech Production* (73–78).

Gick, B. (2002). The use of ultrasound for linguistic phonetic fieldwork. *Journal of the International Phonetic Association*, *32*, 113–121.

Modha, G., Bernhardt, B., Church, R. & Bacsfalvi, P. (2008). Ultrasound in

treatment of /r/: A case study. *International Journal of Language & Communication Disorders, 43*, 323–329.
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.
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*, *42*, 37–44.
The National Joint Committee for the Communication Needs of Persons with Severe Disabilities (1992). Guidelines for meeting the communication needs of persons with severe disabilities. *ASHA, 34*(Suppl. 71), 1–8.

**Appendix A**

*Table 1: Stimuli Used in Sessions 1-6 with Ultrasound*

| Session 1+2 | Session 3 | Session 4 | Session 5 | Session 6 | Session 7 |
|---|---|---|---|---|---|
| arrrrrrr<br>grrrrrrr | crazy<br>pray<br>frizzy<br>drizzy<br>breezy<br>frayed<br>tried<br>stay | zipper<br>feather<br>alligator<br>November<br>grasshopper<br>computer<br>butter<br>October | reek<br>rip<br>rake<br>rep<br>rat<br>rut<br>room<br>rookie<br>wrote<br>raw<br>rot | ear<br>erk<br>farther<br>dork<br>tear<br>horse<br>fort<br>fur<br>park<br>river<br>bird | Car<br>dark<br>flair<br>tear<br>clear<br>tire<br>diary<br>sore<br>mortal<br>sir<br>germs<br>girl<br>barrel |

**Appendix B**

*Table 2: Stimuli Sentences used in Session 8*

Theresa rakes red leaves in October
Randy wrote a letter to a girl
The hairy gorilla ate the fruit salad
The rain poured on the umbrella
Thirteen grasshoppers hurried into the yard

She played her guitar at the bar downstairs
Rex was furious that the raccoons ate his popcorn
The tray of buttercup marshmallows
On Thursday, Lorna threw three red rocks
The thermometer read one hundred and four
She opened her presents and poured champagne
The horse heard a noise in the barn
He was reading the Lord of the Rings
The sky was clear yesterday after the fire
The barrel of beer fell on the ground
The air was warm form the rays of sun
Randy has dreads in his hair
He got rid of the stale hamburger
Chris read his story to her pet spider
She could hear the storm and prayed it wouldn’t rain