2012

Effectiveness of manual gesture treatment on residual /r/ articulation errors

Jessica Lynch

EFFECTIVENESS OF MANUAL GESTURE TREATMENT ON RESIDUAL /r/ ARTICULATION ERRORS

A Thesis

Submitted to the Rangos School of Health Sciences

Duquesne University

In partial fulfillment of the requirements for

the degree of Master of Science in Speech-Language Pathology

By

Jessica E. Lynch

August 2012

EFFECTIVENESS OF MANUAL GESTURE TREATMENT ON RESIDUAL /r/ ARTICULATION ERRORS

By

Jessica E. Lynch

Approved July 10, 2012

____
Heather Leavy Rusiewicz, Ph.D.
Assistant Professor of Speech-Language Pathology
(Committee Chair)

____
Diane L. Williams, Ph.D.
Associate Professor of Speech-Language Pathology
(Committee Member)

____
Caterina F. Staltari, M.A.
Clinical Instructor of Speech-Language Pathology
(Committee Member)

____
Gregory H. Frazer, Ph.D.
Dean, Rangos School of Health Sciences
Professor of Health Sciences

____
Mikael D.Z. Kimelman, Ph.D.
Chair, Speech-Language Pathology
Associate Professor of Speech-Language Pathology

ABSTRACT

EFFECTIVENESS OF MANUAL GESTURE TREATMENT ON RESIDUAL /r/ ARTICULATION ERRORS

By

Jessica E. Lynch

August 2012

Thesis supervised by Heather Leavy Rusiewicz, PhD.

The functional speech sound disorder, American English /r/ articulation errors, presents a unique and confounding clinical challenge as “therapy resistant” residual errors persist into adolescence and adulthood in many cases. Finding paucity of empirical research for /r/ treatment, evidence-based practice (EBP) exploration in motor-related disorders informed clinical practice and research directions. This study investigated the efficacy of “manual mimicry” (a kinesthetic, gestural, and visual cue) in treating intractable /r/ errors in a young adult using a single subject ABAB design. Perceptual accuracy judgments of three types of listeners (experts, graduate clinician, and naïve listeners) indicated a positive treatment effect of manual mimicry cueing on vocalic /r/ sound productions. Electropalatograpy (EPG) outcome measures showed limited

ability to accurately reflect perceptual changes quantitatively. These findings from an exploratory study provide initial evidence that perceptual saliency of /r/ productions may be potentially remediated using a kinesthetic, gestural, and visual cue during treatment.

ACKNOWLEDGEMENT

It's always surprising how life can change one's direction and focus with the slightest nudge. Four years ago I would not have imagined that I would be defending a thesis or could have learned so much. It was possible because of the support and encouragement of my family, friends, and colleagues. I appreciate Duquesne University's SLP program for a thorough and experience- rich education. Thank you: Dean Frazier, for your support of me and the program; Mikael Kimelman, Ph.D., for the structure of the program, practical life advice, and praise when deserved; and the faculty and staff, for the professionalism, kindness, and generosity in every interaction with the students. I am indebted to all of my mentors, specifically: my Duquesne University Clinical Instructors (Lori Marra, Annette Neff, Gary Rentchler, Ph.D., Caterina Staltari, Mikael Kimelman, Ph.D., Sarah Wallace, Ph.D., Heather Rusiewicz, Ph.D., and Diane Williams, Ph.D.) for sharing their clinical knowledge; my externship placement supervisors (Miriam Perriello, Carol Robson, Audrey Rozell, Pamela Scheffler, Liza Linn, Diane Ziggas, and Cathy Brumbaugh) for their understanding, support, mentorship, and friendship; and all of my aforementioned professors (including Rob Masterson and Susan Snyder). I am very grateful to Diane Williams, Ph.D., Sarah Wallace, Ph.D., Heather Rusiewicz, Ph.D., Yang Chen, Ph.D., and Leah Helou, Ph.D. for sharing their love and passion for research, imbuing me with a desire to find answers and contribute to the field, as they have. An acknowledgment to Tara McAllister Byun, Ph.D. and Elaine Russo Hitchcock, Ph.D. from Montclair University for presenting a session at the ASHA convention in Philadelphia in 2011 about treating /r/, which allowed me to realize the clinical demand for treatment knowledge in this area when the sessions were standing room only, needed to be repeated, and still had to turn away numerous SLPs.

Thank you to my thesis committee, Heather Rusiewicz, Ph.D., Diane Williams, Ph.D., and Caterina Staltari for the thoughtful feedback, direction, flexibility, and insight. Diane, I appreciate you believing in my abilities, challenging me to reach further, and helping me expect more from myself. Caterina, I am amazed by your wealth of clinical expertise, skill, knowledge, and acumen and am grateful for your willingness to share theoretical rationale and tips - you are magic. Heather, a huge thank you for endless hours of theoretical discussions and proofreading, professional advice, guidance toward interesting research, support, and friendship, but most of all for encouraging me never to stop even when it seemed impossible. Thank you, Class of 2012 for keeping your humor and accepting my propensity to intellectualize. None of this could have occurred without the love of my family and dearest friends who make me so grateful. Tony, you commiserate and debate with me and keep things fun. Jana, your friendship, unconditional love, ever present ear, and encouragement are invaluable blessings. For my siblings, and especially my nephew, Julian, for understanding when I had school and loving me even when they missed me. Mom, your constant desire to see me grow into the woman that you already know I am, subtle nudges toward success, unending ear and shoulder, willingness to sacrifice, and undying love are more than I can ask. Dad, for always providing a calm intelligence with a willingness to discuss any topic, and showing interest in endless problem solving and boring technical discussions. Grandpap, the stability and support that you have given our family is a valued gift without question. Rowena, Donald Sr., and Donald Jr., for endless, selfless, support, assistance, stability,

strength, faith, love, pride, and constant nurturing of my talents and dreams. Palmer, you stuck out the struggles and transitions, and loved me unconditionally even when it was hard. Thank you for always listening to my grad school talks (and even for learning SLP lingo along the way).

TABLE OF CONTENTS

LIST OF FIGURES

Introduction

Articulation disorders are classified by numerous terms based on etiology and presentation. Articulation disorders involve errors in speech sound production including substitutions, distortions, and omissions of sounds (Peña-Brooks & Hedge, 2007, p. 352). Functional or idiopathic speech sound disorders (SSD) are ones that have no known etiology but are attributed to the motor speech breakdown in the execution level of the speech system. The speech sound error this paper specifically addresses is the residual, persistent, long-standing, or therapy-resistant error of the American English /r/ phoneme (Byun & Hitchcock, 2011).

The /r/ phoneme is classified as a liquid, lingua-palatal sound, and it possesses the following characteristics: voiced, vocalic, consonantal, coronal, round, continuant, sonorant, syllabic, back consonant, approximant, and rhotic (Peña-Brooks & Hedge, 2007, p. 80-81). The /r/ sound is the fourth most frequently occurring sound for English consonants, preceded only by /n/, /t/, and /s/ (Shriberg & Kwiatkowski, 1983, as cited in Peña-Brooks & Hedge, 2007, p. 7.). It is one of the later acquired sounds for children (Sander, 1972). According to Sander (1972), the average age at which 90% of all children are customarily producing the /r/ sound is age six. It is one of the most common speech sound errors in school-age children (St. Louis, Ruscello, & Lundeen, 1992, as cited in Shuster, Ruscello, & Toth, 1995, p. 37). The distortion of /r/ is considered a residual articulation error, due to the tendency of /r/ articulation errors to be lingering distortions of atypical adolescent and adult productions. The distortion of the American English phoneme /r/ has an obvious impact on the ability of individuals with this error to communicate because of the frequency with which this sound occurs in American English.

Statement of the Problem

When the distortion of /r/ persists through adolescence and into adulthood, it has a psychosocial implication for some individuals. The speech disorder is easily identified by peers and generally judged negatively, which has far reaching implications into adulthood (Crowe Hall, 1991; Silverman & Paulus, 1989). Felesenfeld and colleagues (1994) suggested that the repercussions of the social and academic participation limitations resulting from a speech disorder may be life-long.

The correction of the /r/ distortion is one of the more challenging intervention goals that speech-language pathologists encounter during intervention (Bernthal & Bankson, 1993; Clark, Schwarz, & Blakely, 1993; Secord, 1981; Shriberg, 1975, 1980; Shuster et al., 1995). In Ruscello's (1995) survey of SLPs in public schools, /r/ was the most commonly reported sound that the children failed to obtain and 91% of SLPs responding reported having a client who did not acquire this speech sound following therapy, with 40% reporting discharge of clients prior to effective remediation. Reasons for difficultly in remediating the /r/ sound abound in research literature from assertions regarding within and between speaker variations in typical /r/ production to acoustic differences (Delattre & Freeman, 1968; Kent & Read, 1992; Ohde &Sharf, 1992; Shriberg & Kent, 1982; Zwadski & Kuehn, 1980, as cited in Shuster et al., 1995).

Intervention challenges are compounded by the lack of visual input to the client for direct modeling of placement. In addition, there are numerous well-documented tongue shape variations, including the most common extreme shapes for typical /r/ production: retroflex "produced with a raised tongue tip and lowered [dorsum]," humped or bunched "produced with a lowered tongue tip and raised dorsum" (Zhou, Epsy-Wilson, Boyce, Tiede, Holland, & Choe, 2008, p. 4466), and lateral rhotic (Haynes & Pindzola, 2008, p. 75). Research to date has not found consistent patterns of tongue shape, vocal tract constriction, or formant frequencies in typical speakers (Zhou et al., 2008; Alwan, Narayanan, & Haker, 1996). Twist and colleagues (2007) discovered that listeners cannot detect the difference between the two most common /r/ tongue configurations, retroflex and bunched /r/ (as cited in Zhou et al., 2008). Vocalic variations of the /r / sound (i.e., /εr/, /ir/, /or/, /Ɪr/, or /αr/), the stressed (hooked schwa, schwa r, or /ɚ/), and unstressed /ɝ/ also complicate the learning and generalization of correct articulation (Peña-Brooks & Hedge, 2007, p. 285). These complications are based on co-articulation of surrounding phonemes, the position in the word, rhotic "coloring" of vowels, and rhotic vowel variability (Haynes, 2008, p. 171). Christine Ristuccia (2002) defines 21 different types of /r/ as the following:

/air/ Initial	/air/ Medial	/air/ Final
/ar/ Initial	/ar/ Medial	/ar/ Final
/or/ Initial	/or/ Medial	/or/ Final
/ear/ Initial	/ear/ Medial	/ear/ Final
/ire/ Initial	/ire/ Medial	/ire/ Final
/er/ Initial	/er/ Medial Stressed	/er/ Final
Prevocalic /r/	/er/ Medial Unstressed	/rl/ Medial & Final

Figure 1. Christine Ristuccia (2002). 21 different types of /r/.

History of articulation therapy techniques for /r/. Therapy techniques for /r/ articulation errors date back to the early 1900s; however, very limited empirical research data exists with regards to remediating /r/ errors. Gibbon and Patterson (2006) stated that in speech therapy research, /r/ is one of the most neglected areas. Therefore, clinical practice of /r/ remediation is primarily based on expert opinion and clinical judgment levels of evidence. A review of the history of /r/ specific articulation therapy follows.

Traditional auditory feedback cueing involves approximating the /r/ sound by starting with /ɝ/ and providing placement cues (Mowrer, 1975; Shriberg, 1975, 1980). Sound evocation programs are also provided for /r/. A milestone in the development of articulation therapy occurred when Spriesterbach and Curtis (1951) contended that both /r/ and /s/ misarticulations could be produced in specific contexts and stimulated using facilitative phonetic contexts. Van Riper and Irwin (1958) concurred by claiming that certain phonetic environments facilitate correct production (Peña-Brooks & Hedge, 2007, p. 323). For example, if a child cannot usually say the /r/ sound, but in certain words such as "spring" the sound is accurate, then the "spr" context is a facilitative phonetic environment for that child. Out of this concept, Eugene McDonald developed his *Deep*

Test of Articulation (McDonald, 1964). Shaping, phonetic placements, and facilitative contexts are used presently in clinical settings for /r/ treatment.

Tactile feedback. Oral motor exercises were also a popular therapy choice but with empirically refuted efficacy (Peña-Brooks & Hedge, 2007, p. 440). Shriberg (1980) created an intervention technique for intractable /r/ errors, in which a bite stick is used to stabilize the jaw to prevent excessive movement, and then direct models are provided for imitation and perceptual judgment of correctness (Shuster et al., 1995). This program also uses auditory feedback and tactile or kinesthetic cues. A randomized group experiment indicated that a removable /r/ appliance that provided tactile feedback from a prosthetic device was effective in remediating /r/ (Clark et al., 1993). Therefore, tactile feedback has proven useful to elicit the more challenging /r/ sound.

Instrumentation for remediation of /r/. In the mid-1990s, technological advancements made it possible to use instrumental equipment to treat /r/ errors. Although limited empirical research exists for /r/ remediation using traditional therapy approaches, numerous studies of both assessment and remediation of /r/ using instrumentation have been conducted and provide empirical evidence of treatment efficacy. Instruments are mainly used for measuring tongue (and other articulators) placement and movement during speech samples. Instrumental research has been proven as an effective treatment tool to remediate /r/ errors through the use of visual feedback, or more specifically visual biofeedback.

Many instruments have been used to measure and to provide visual feedback during intervention including: electropalatography and glossometry; x-ray, ultrasound, and magnetic resonance imaging (MRI) imaging techniques; electromagnetic articulography (EMA) and electromagnetic midsagittal articulography (EMMA) motion tracking instruments; and optoelectric optical motion systems (Gibbon, 2008). Gibbon (2008) stated, "Growing literature showing the effectiveness of using visual feedback…to improve intelligibility is likely to further promote their clinical use" (p. 326).

Instrumentation equipment can provide much needed visual feedback to clients with articulation errors, especially for production of sounds like the /r/ phoneme that are produced toward the back of the oral cavity and hence intrinsically lack visual feedback. The effectiveness of visual feedback for treatment of articulation and phonological disorders has been demonstrated in the literature (Gibbon, 1999).

According to Rvachew and Nowak (2001), current techniques to remediate /r/ errors, which were the most difficult for their participants to learn, included "auditory-perceptual cues (Rvachew, 1994), visual-feedback (Shuster, Ruscello, & Toth, 1995), or a minimal-pairs (Gierut, 1989) approach to treatment" (p. 621). Visual feedback appears to provide some consistently promising data in articulation remediation. One case study was found to have utilized ultrasound for /r/ therapy (Adler-Bock, Bernhardt, Gick, & Bacsfalvi, 2007). Another case study used the visual biofeedback of a spectrogram to correct /r/ errors (Shuster, Ruscello, & Toth, 1995). Other therapeutic strategies include: visual feedback from spectrograms, biofeedback, and oral modifications (Guilford & Hnath-Chisholm, 1991; Hardcastle, Gibbon, & Jones, 1991; Netsell & Cleeland, 1973; Resiberg, 1968; Ruscello, Cartwright, Haines, & Shuster, 1993; Shuster, Ruscello, & Smith, 1992; Wolfe & Irwin, 1975; Gibbon, Hardcastle, & Suzuki, 1991; Clark, Schwarz, & Blakely, 1993, as cited in Shuster et al., 1995).

Visual biofeedback using electropalatography has been demonstrated to be efficacious with various types of speech production disorders including articulation and phonological disorders (Gibbon, 1999; Hardcastle, Gibbons, & Jones, 1991; Dagenais, 1995; Gibbon, Stewart, Hardcastle, & Crampin, 1999; Gibbon 2008; Howard, 2007). Limiting factors in the use of this technology are cost and accessibility. Equipment and instrumental remediation are not currently available in most clinical settings.

General articulation therapies. When searching for EBP treatments for /r/, clinicians may tend to gravitate toward traditional articulation therapy, even if it is not specifically focused on the characteristic challenges of /r/ errors. Treatments specific to /r/ errors (non-instrumental treatments, specifically) are limited; therefore, general articulation therapy has been traditionally utilized to remediate /r/ difficulties. One articulation therapy that remains dominant in current clinical practice was developed first by Stichfield and Young (1938), and then further developed by Van Riper (1939) who advocated use of a five phase progression from sensory perceptual training to maintenance. Van Riper and Erickson (1996) updated the model in the mid- 1990s by adding a visual staircase to symbolize progression through the steps (Peña-Brooks & Hedge, 2007, p. 399). There is controversy surrounding efficacy of certain elements of traditional articulation therapies. One controversy questions the efficacy of Van Riper’s traditional therapy given its omission of phonological pattern analyses currently used in differential diagnosis between articulation and phonological disorders. Additionally, there is controversy over the effectiveness and necessity of the sensory perceptual training element of Van Riper’s treatment. Otherwise, traditional therapy remains clinically dominant, but “no systematic experimental evaluation of the total approach and of the different elements of the approach" has been conducted (Peña-Brooks & Hedge, 2007, p. 407). Similarly, no controlled research of the effectiveness of facilitative contexts as compared to other therapy approaches exists. Controversy over use of stimulable sounds for shaping ease versus more complex or difficult sounds for enhanced generalization effects still persists (Peña-Brooks & Hedge, 2007, p. 411).

Scripture and Jackson (1927) advocated the use of phonetic placement techniques using a behavioral process later coined as "shaping." This approach promoted accurate productions through direct teaching methods that ensured that the child knew how to produce the sound correctly by explaining articulator positioning, modification of air stream for correct manner of production, and voicing elements (Peña-Brooks & Hedge, 2007, p. 469).

As synthesized by Peña-Brooks and Hedge (2007), the intervention techniques supported by experimental research include: phonetic placement (p. 469); successive approximation and sound shaping (p. 470); modeling (Creaghead et al., 1989; Hedge, 1998); imitation (Hedge, 1998); drawing attention to the kinesthetic properties of a sound by focusing on how sound production feels (Bankson &Bernthal, 2004); vocal emphasis techniques used by clinicians to draw attention to target sounds while providing instructions verbally (Hedge, 1998); and, prompts (Peña-Brooks & Hedge, 2007, p. 471-472). These intervention techniques are described herein. Modeling is when the clinician produces and models a sound the client is expected to make and imitation is the client's response to the clinician's model (Hedge, 1998). Vocal emphasis techniques consist of increasing intensity and length of the sound coupled with frequent modeling (Hedge, 1998). Verbal instructions are stimuli given verbally about how to produce a sound (Hedge, 1998). Prompts can be verbal or non-verbal and are used to elicit the occurrence of the target production and to increase the probability of accuracy. They are "[thought of] as "hints" or "cues" to draw the response from the [client]" (Peña-Brooks & Hedge, 2007, p. 472). A verbal prompt (e.g., vocal emphasis), traditionally called verbal cues or auditory stimulation, are used as reminders of therapy training (e.g., "Remember, your tongue stays inside of your mouth for our sound"). Non-verbal or physical prompts (i.e., visual cues or visual stimulation) are used often in articulation therapy (Peña-Brooks & Hedge, 2007, p. 472-474).

Multi-sensory articulation treatments. Articulation therapy often incorporates multi-modality cueing. McDonald added a new approach to articulation treatment by focusing on the feel of articulator placement. He focused on the awareness aspect or what is now termed tactile-kinesthetic cueing. The sensory motor approach (SMA) developed by McDonald (1964) had the primary goal of increasing auditory, tactile, and proprioceptive awareness of motor patterns in speech sound productions using motor production tasks in facilitative phonetic contexts (Peña-Brooks & Hedge, 2007, p. 407). These strategies are similar to PROMPT (the prompts for restructuring oral muscular phonetic targets) treatment used effectively in therapy for childhood apraxia of speech (CAS) (Freed, Marshall, & Frazier, 1997). Shrine and Proust (1982) made two adaptations to the SMA by extending the number of phases to two and adding more behavioral therapy treatment paradigm elements. Their adaptation of SMA is a current remediation practice (Peña-Brooks & Hedge, 2007, p. 407).

Weiss, Gordon, and Lillywhite (1987) built on the concept of sound shaping techniques and advocated using phonetic placement instruments including: tongue blades, breath instruments, and graphic records (e.g., spectrograms). They also suggested shaping sound by using the following techniques: diagrams, pictures, drawings of articulators, palatograms, observation in mirror, clinician manipulation of client's articulators, verbal description and instruction, and feeling the breath stream and laryngeal vibration with hand. Hedge (1998) described the clinician's physical manipulation of the client's articulation as manual guidance, which is also described as tactile-kinesthetic cueing or tactile- kinesthetic stimulation (Peña-Brooks & Hedge, 2007, p. 469-470). Secord (1989) suggested that anatomic descriptions and more sophisticated terms describing auditory, visual, and tactile-kinesthetic properties should be used with older children and adults. This multimodality strategy continues to be one of the dominant therapy techniques used in current practice.

Empirical Support

Empirically supported aspects of articulation therapies. It is important to know what treatments are backed by empirical research to ensure that best practices for Evidence Based Practice (EBP) are employed by clinicians; treatment time is valuable and must be used wisely. Clinically relevant therapeutic methods for the remediation for /r/ are necessary, especially given the frequent occurrence of the /r/ distortion on the school-based SLP caseload and the impact on individuals with intractable /r/ errors.

The specific elements of articulation treatment known to be effective include: "modeling, systematic positive reinforcement, and corrective feedback for incorrect productions." The elements of articulation treatment thought to be effective include:

"repeated practice of speech sound production, varied phonetic and linguistic contexts, [extending skills to] conversational speech [in] natural contexts, and teaching self-monitoring" (Peña-Brooks & Hedge, 2007, p. 407).

Empirically supported behavior therapy. Behavioral therapy techniques are other treatment paradigms that have been adopted clinically due to the body of empirical research supporting their use for effectively changing behavior. As formal therapy programs have systematically included advancements from the field of psychology, behavioral therapy techniques have been commonly applied to speech therapy practice. A few examples of this adaptation to speech therapy follow. The Paired-Stimuli Approach by Irwin and Weston (1971; 1975) was a sequenced and highly structured program based on operant learning, behavioral therapy contingencies, and stimulus-response generation capitalizing on the facilitative effects of key words, or contexts in which the target sound can be produced correctly. Programmed conditioning for articulation (PCA) by Baker and Ryan (1971) was a theoretical and behavioral treatment that used programmed instruction and learning concepts based in operant conditioning and stimulus-response-consequence contingency paradigm developed by applied behavioral psychologists (Peña-Brooks & Hedge, 2007, p. 421-422).

Controlled studies are lacking for the paired-stimuli approach; although, specific elements of these approaches have efficacy evidence. PCA is based on the effective procedures of: modeling, positive reinforcement, corrective feedback, and systematically shaping complex skills from simpler skills (Peña-Brooks & Hedge, 2007, p. 425). Data from the authors of these therapies suggest improvement from their application;

nevertheless, controlled treatment studies are needed to establish the effectiveness of these therapy paradigms.

Examining /r/ articulation therapies. Aspects of traditional articulation therapy, behavior therapy, and visual biofeedback have demonstrated efficacy for the treatment of /r/. However, the clinical challenge of remediating /r/ remains. Clinicians who have used these traditional articulation techniques coupled with behavior therapy have reported limited success remediating /r/ speech sound errors. As previously stated, most school-based SLPs do not have access to the instrumental visual biofeedback equipment preventing them from utilizing this approach even if it has been shown to be efficacious. Because access to equipment is not always reasonable, clinicians are seeking novel treatment approaches to /r/ errors. Examining the literature on treatment research for similar populations may lead to potential therapeutic options for /r/ remediation. Research in motor-speech disordered populations (CAS) may lend insight into treatment options for /r/. Because the literature for /r/ supports the use of visual feedback as effective treatment, exploring the visual feedback treatment research for other speech sound disorders or CAS may lead to therapy techniques that could also be used successfully for /r/ errors.

Research supported therapies for other SSDs that may be useful. Various intervention programs targeting CAS may contain useful strategies for articulation disorders including: the work of DeThorne and colleagues (2009) for CAS; integral stimulation therapy (Rosenbek, Lemme, Ahren, Harris, & Wertz, 1973; Strand & Debertine, 2000, as cited in Edeal & Gildersleeve-Neumann, 2011); and dynamic temporal and tactile cueing (DTTC; Strand, Stoeckel, & Baas, 2006; Berman, Garcia, &

Bauman-Waengler, 2007; Daniel, 2009; Jakielski, Webb, & Gilbraith, 2006; Jensen & Glidersleeve-Neumann, 2005, as cited in Edeal & Gildersleeve-Neumann, 2011). The effectiveness of integral stimulation therapy and its pediatric version DTTC (Strand et al., 2006 as cited in Edeal & Gildersleeve-Neumann, 2011) is supported by single-subject design studies (Strand & Debertine, 2000, as cited in Edeal & Gildersleeve-Neumann, 2011). Integral stimulation therapy incorporates principals of motor learning into a motor-based hierarchical speech therapy. Hierarchies are used for both target selection and level of cueing support. DTTC adds multimodal cueing techniques (auditory, visual, and tactile) simultaneously for maximum cueing support that can be modified or reduced according to hierarchical needs (Edeal & Gildersleeve-Neumann, 2011). A similar multi-modality cueing technique is the Signed Target Phoneme (STP) treatment. Shelton and Graves (1985) conducted a case study in apraxia of speech using STP treatment as a visual approach, in which hand shape and verbal representation of the sound were used simultaneously.

In a general review of motor learning intervention techniques, DeThorne and colleagues (2009) focused specifically on interventions for speech sound development early in life with clients who are not responding to imitation strategies. According to DeThorne and colleagues some research based intervention strategies that may be used when imitation is not effective include: providing access to augmentative and alternative communication (AAC), minimizing pressure to speak, imitating the child, utilizing exaggerated intonation and slowed tempo, augmenting [multimodality] feedback, focusing on function with articulator movements, and avoiding emphasis on non-speech articulator movements.

The above mentioned therapy strategies are supported by sound empirical evidence when used for SSDs and CAS intervention. The techniques referenced above use multimodality feedback (i.e., augmented auditory, visual, tactile, and proprioceptive feedback) to elicit correct sound production. Use of multimodalities is grounded in motor learning theory and appears to be vital to speech and language learning because all learners do not respond to the same methods. Using best practices that incorporate all modalities will increase the likelihood that the lesson is learned and retained. Intervention practice using multimodalities is becoming more widely accepted and receiving acclaim for its effectiveness. DeThorne and colleagues (2009) emphasize a "task-dependent view [that] stresses the importance [of forming an internal model of the target] and emphasizes the goal or purpose of the motor task as critical to its generalization" (p. 140). A client's internal model can best be facilitated through preferred modalities.

More research is needed in the areas of motor speech learning, speech production, and interventions for speech sound disorders. However, there is already valuable information from motor speech learning theory and related theoretically based interventions for speech sound disorders that can serve as a starting point for EBP for the correction of residual /r/ errors.

Exploring Non-Verbal or Physical Prompts

According to Peña-Brooks and Hedge (2007):

[Non-verbal or physical prompts] can be thought of as physical signs and gestures that may help the client visualize correct production of the target sound. In non-verbal prompts, the clinician may physically prompt correct production of a target sound by using his or her hands for demonstration. [It] comes naturally to clinicians, and it can be a powerful facilitative technique for sound establishment if used appropriately and in conjunction with other methods such as verbal instruction and modeling (p. 473).

Prompting clients non-verbally can be done in many ways, including the use of visual feedback. Non-verbal prompts may be preferred by practicing clinicians due to limitations that the instrumental visual feedback may have, such as cost, accessibility, and lack of generalization to more natural environments. A few examples of non-verbal prompts follow. Mowrer (1989) included diagrams and a written stimulus as physical or non-verbal prompts. It should be noted that physical prompts are not the same as manual guidance, because, unlike manual guidance, the clinician demonstrates, but does not have direct physical contact with the client (Mowrer, 1989; Hedge, 1998). In addition, non-verbal informative feedback is sometimes used as secondary reinforcement (e.g., graphs and charts, sometimes on a computer monitor) (Peña-Brooks & Hedge, 2007, p. 479).

Visual cues are used frequently by clinicians (i.e., pointing to your mouth to gain client focus, modeling or demonstrating articulator placement, and hand gestures imitating tongue movement). Other examples of visual cues include: pointing to salient articulatory placement, using hands to mimic articulator movement and timing (herein referred to as “manual mimicry”), or a gesture used as a reminder to the client to self-cue previously trained skills. Using non-verbal prompts, especially those with visual elements needs to be explored further as a potential source for effective /r/ therapy treatment.

Manual movement cues. As mentioned above, in the treatment of motor speech disorders, namely apraxia of speech, multi-modality cueing (using extensive tactile-kinesthetic, visual, and non-verbal cues in addition to what is called "gestural input" and "manual symbols") is common clinical practice (Square, 1999). These techniques have been used to pace speech and give cues regarding vocal tract configurations, and real time movement of articulators (Square, 1999). There are some treatments which may or may not be considered gestural, that involve using hands to elicit more accurate speech production. Carahaly (2012) described a program using multisensory cues (auditory, visual, proprioceptive, gestural, and tactile cues) to treat CAS called The Speech EZ program. Some of the more symbolic "gestural" treatments include: adapted cueing therapy (ACT), Jordan's gestures, signed target phoneme (STP) therapy, and cued speech (Square, 1999, p. 175-176).

ACT and STP. ACT uses hands positioned near the clinician's face to visually demonstrate to the client the articulator movement and manner of production while providing an auditory model as well. It was developed by Klick (1984, 1994), loosely based in American Sign Language (ASL), as a "gestural presentation of speech in motion" (Square, 1999, p. 175). There are no studies that provide empirical evidence for ACT and the concept as explained to clients is criticized for being complex. The cognitive load for both the client and the clinician for this treatment may affect the treatment efficacy due to challenges in teaching, learning, and implementing ACT. STP developed by Shelton and Graves (1985) is a simplified version of ACT in that only the target phoneme is demonstrated by the clinician within in the word or utterance (Square, 1999).

Jordan's gestures. Similar to STP, Jordan's gestures (Jordan 1988, 1991) are visual and symbolic of individual phonemes but represent: vocal tract configuration, transitional oral movements, articulatory contact points, and voicing elements (Square, 1999). Jordan's gestures are the most similar gestural symbols to the hand gestures proposed in this project. In that gesture system, the /r/ phoneme is created by extending fingers slightly and touching the thumb to them. The kinesthetic movement associated with the gesture is moving the thumb back along the inside of the fingers and "attention is called to the 'feeling' of tension at the base of the thumb with the palm" (Square, 1999, p. 249).

Figure 2. An approximate example of the Jordan's gesture for the /r/ phoneme.

In a study using Jordan's gestures with children who had CAS from 1992, Hall and Jordan (as cited in Square, 1999, p. 254-259) assert that:

> [The participants] may have found helpful the gestural information about the point of constriction of articulator placement and movement, which facilitated their correct production of the /tʃ/, /s/, /ʃ/, and /r/ ["ch", "s", "sh", and "r", respectively]. The children were observed to spontaneously use the gestures when making production attempts during sessions, thus seemingly benefitting from self-cueing aspects of the system [Jordan's gestures]. It was further noted

> that the use of self-cueing decreased as the oral production skills with the targeted phonemes became more consistent. Thus, the gestures may have enhanced the children's overall sensory awareness for articulatory placement, movement, and voicing by use of visual, tactile, and kinesthetic, as well as auditory input (p. 254-259).

One aspect of note in Hall and Jordan's (1992) study is that the hand gestures were used by the clinician during training, but then the client spontaneously mimicked the hand movements to self-cue for accuracy. The clinician demonstrating the manual movement is a form of externally generated cue; however, the client performing the "manual mimicry" hand movement serves as an internal cue that can be spontaneously incorporated, assist in generalization and carry-over, and serve as an internal reminder of correct production, as discussed in the section on principals of motor learning.

Research on gesture use as clinical cueing. Examining gestures used as clinical prompts during treatment may provide useful information that can be applied to remediation for /r/ distortions. This burgeoning research topic from many fields of study has focused on both the gestures made by clinicians as prompts and those made by clients as self-prompts. In speech-language pathology, research on gesture has been conducted with numerous populations with varied research foci. Research potentially relevant to this topic includes: listener understanding, listener reaction time, and memory processes (Riseborough, 1981, as cited in Garcia, Cannito, & Dagenais, 2000, p. 109), increased learning in teaching contexts (Cook & Goldin-Meadow, 2006, p. 211), increased strength of memory encoding (Rugg & Curran, 2007, as cited in Kelly, 2008, p. 13), decreased work load on Broca's area for processing meaning (Skipper, Goldin-Meadow, Nusbaum,

& Small, 2007, as cited in Kelly, 2008, p. 5), and the ability to mimic non-verbal behaviors and gestures that we see, both goal directed and not (Chartrand & Bargh, 1999; Meltzoff & Moore, 1977; Meltzoff 1988; Bekkering, Wohlschleager, & Gattis, 2000; Carpenter, Call, & Tomasello, 2005, as cited in Cook & Goldin-Meadow, 2006, p. 212). According to Cook and Goldin-Meadow (2006):

> As another possibility gesture uses the body to do its representational work, and these embodied representations might promote learning. There is increasing evidence that embodied forms of representation are involved in cognitive processes, including working memory (Wilson, 2001), action memory (Englekamp, 1998; Nilsson et al., 2000), mental imagery (Jeannerod, 1995; Kosslyn, 1994)….Gesture, as an embodied representational format, could preferentially engage [any of these] systems in contributing to learning (p. 228).

Findings indicate that gesture is present in speakers when listeners are not present, which indicates that the speaker uses gestures to self-cue or assist his/her speech efforts and not necessarily for the sole benefit of the listener. For example, individuals who were congenitally blind gestured when speaking to other blind children, indicating that speakers use gesture as an internal communicative function (Cohen, 1977; Iverson & Goldin-Meadow, 1998; Wolff & Gustein, 1972; Bull, 1983, as cited in Garcia et al., 2000, p. 109). Therefore, gestures may promote motor learning and may be able to be used as a method of self-cueing.

Additional research also indicates that gestured items, both verbally and visuo-spatially, may be remembered more than words not gestured or words given in verbal form only (Cohen, 1981; Engelkamp & Krumnacker, 1980; Saltz & Donnenwerth-Nloan, 1981; Bahrick & Bouche, 1968; Durso & Johnson, 1980; Paivio & Csapo, 1973; Goldwin-Meadow, Nusbaum, Kelly, &Wagner, 2001; Clark & Paivio, 1991; Wagner, Nusbaum, & Goldin-Meadow, 2004, as cited in Kelly, 2008, p. 11). Therefore, gesture may improve a client's memory of the gestured items, promoting generalization of therapy targets to other situations.

Gestures may help individuals with motor-based speech disorders increase articulatory accuracy. Garcia, Cannito, and Dagenais (2000) discussed gesticulations as a compensatory strategy for people with dysarthria; in addition they noted that beat gestures may "contribute to more precise articulation and natural sounding speech" (p. 113). It is also interesting to note that in another study (Garcia and Cannito, 1996) the intelligibility of a person with dysarthria increased when using gestures in an auditory only condition in which the listener heard but could not see the participant's gesture use. Consistent with the work with blind children, this suggests that the individuals with dysarthria were using the gesture for self-cueing. If gesture can indeed increase production accuracy for individuals with dysarthria, it may be valuable in remediating residual articulation disorders as well.

Given the research regarding the use of gesture mentioned above indicating 1) that we remember what we gesture more readily 2) that gesture use may increase speech intelligibility, and, 3) that gesture provides a form of visual cueing, this project proposes to examine gesture use by the clinician to improve the client's speech production accuracy. More specifically, the current study proposes exploring use of a hand movement (manual mimicry) to increase intelligibility (in /r/ articulation disordered populations). This hand movement is a motorically analogous movement to the speech mechanism, potentially tapping into the cognitive representation of the required behavior as well as a coordinated entrained movement across systems.

Discussion of Rationale for Studying this Clinically Based /r/ Treatment

This study was designed to examine the use of a hand movement as an accessible tool for clinicians in need of treatment options for therapy resistant /r/ errors for the clients they serve thereby diminishing psychosocial impact. Developing methods for /r/ error correction is warranted because of the distortion's frequent occurrence, its recognized remediation difficulty, and the psychosocial implications for some individuals when this distortion persists through adolescence and into adulthood. Furthermore, even though the use of instrumentation can be useful in remediation of /r/, it can be expensive and inaccessible in the school systems, where the majority of /r/ treatment is conducted; therefore, it is important to have effective treatments available that do not involve costly equipment. The therapy options discussed above included few /r/ specific options: traditional therapy with limited evidence; visual instrumental therapies with evidence but high costs; non-specific articulation therapies with limited evidence; multisensory therapies with sound evidence for other populations; empirically supported behavioral therapy and aspects of articulation therapy; and promising work in non-verbal, physical, and gesture cues.

Gesture used by an individual has been reported to increase: mental imagery, working memory, self-cueing abilities, proprioceptive awareness, internalization and enhancement of learning, and intelligibility and naturalness, while decreasing the cognitive load. In addition gestural use and multimodality cueing have been shown to be potentially efficacious, particularly with disorders that exhibit difficulty with articulation (namely CAS and dysarthria). Therefore, it appears that manual mimicry cueing may generate a significant effect during the production of the exigent phoneme /r/ in clients with residual /r/ errors. This study examined the efficacy of “manual mimicry” for clinical use in intervention with the challenging /r/ articulation distortion.

Manual mimicry. Considering the success of multimodality cueing reported in the literature, especially when combined with principles of motor learning it warrants further examination in the area of articulation, namely /r/ errors. Research supports the efficacy of: visual feedback; aspects of articulation therapy (i.e., modeling, placement cues); behavioral therapy and PML (i.e., reinforcement, knowledge of results); multisensory therapy (i.e., kinesthetic cue); and nonverbal, physical and gestural prompts (manual mimicry) that could be incorporated into a therapeutic approach to intractable /r/. All of these specific aspects of what we know to be effective at this point could be incorporated into one simple hand gesture (manual mimicry).

This manual mimicry gesture theoretically provides a form of visual (pseudo-biofeedback) by allowing the clinician to approximate the client’s tongue movement (relative to high and low and front and back); timing (assisting with motor planning and execution of the motion of the tongue during speech sounds); shape and placement (tongue configuration within the oral cavity); tension (tense vs. lax can be easily demonstrated with the hand); and simple co-articulations (visual demonstration of tongue movement from a vowel [using the vowel quadrilateral model] to /r/ within the oral cavity). All of this information can be conveyed through the motion of the hand next to the clinician's cheek to cue and feedback can be provided without complexity of excessive verbal explanations. The ultimate goal is to eliminate and simplify the cognitive load on the client while maximizing feedback through efficiency of using manual mimicry.

Manual mimicry as an independent variable. Research has demonstrated the effectiveness of visual biofeedback for /r/ remediation; however, considering the cost and availability of EPG systems (especially to the school SLP), it is vital to find a more cost effective treatment option with comparable EBP and clinical effectiveness. All of the above factors contributed to the decision to use manual mimicry as the independent variable in this study.

Dependent variables. The dependent variables were also selected based on functional impact. The "gold standard" for any articulation therapy success is whether the error can be detected by people in everyday natural contexts and environments; therefore, listener perception of whether the error is correct or incorrect was used as an indication of therapeutic success. Three types of listeners provided data on perceptual accuracy. First, the treating clinician scored perceptions of accuracy (correct/incorrect) on-line and again under a second listening condition with headphones and an audio only wave file recording. Second, two expert listeners scored perceptual accuracy (correct/incorrect) using randomized CD recordings. These two experts also provided phonetic transcriptions of a portion of the recordings. Third, twenty-eight naïve listeners scored perceptual accuracy (correct/incorrect) using randomized CD recordings perception. The final dependent variable incorporated the use of the electropalatography equipment as an outcome measure, instead of as a form of biofeedback treatment. EPG has been reported to be effective in therapy for correcting speech errors; however, to our knowledge it has not been used solely as a dependent variable or outcome measure. This study also contributes valuable information about the use of EPG as an outcome measure of treatment efficacy.

Specific Aim, Experimental Question, and Hypothesis

Specific Aim

To assess the influence of manual cueing (i.e., manual mimicry) treatment on (i) listener perceptual accuracy judgments of a clinician, experts, and naïve listeners (ii) expert clinician transcriptions, and (iii) EPG tongue to palate configuration during vocalic /r/ production probes.

Experimental Question and Hypothesis

1) **Does manual cueing (i.e., manual mimicry treatment) have a significant effect on the accuracy (auditory perceptual and EPG measure) of vocalic /r/ speech production in a young adult with long standing residual /r/ articulation error as compared to a no treatment condition?**

H_0: There is no significant difference between the accuracy of vocalic /r/ productions produced during manual mimicry treatment and the vocalic /r/ productions produced during treatment withdraw phases.

Specific research questions guiding dependent variable selection. Three main questions guided this study**:**

1. Does manual cueing affect /r/ production as measured by electropalatography measures of percentage of accurate palate to tongue contacts?

2. Does manual cueing affect /r/ production as measured by expert listeners' judgments of vocalic /r/ production accuracy?

3. Does manual cueing affect /r/ production as measured by naïve listeners' judgments of vocalic /r/ production accuracy?

Rationale

The theoretical framework of speech sensorimotor control is the basis of our research treatment for an articulation disorder of motor execution; by increasing sensory feedback we enhance the motor plan. The theory at work in this study posits that maximizing sensory feedback will create more accurate articulatory positions. Sensory feedback included 1) visual feedback to augment understanding of motoric movements and 2) tactile-kinesthetic feedback to entrain movements within the participant and (less importantly) between the participant and the clinician)

This project also sought to build upon the finding that visual instrumental feedback is promising for /r/ remediation. Therefore, this project examined how the elements of visual feedback provided by instrumentation with demonstrated effectiveness can be adapted into an accessible therapeutic technique for school SLPs.

Literature Review

Lack of empirically supported /r/ and articulation treatments. Although the therapies discussed in an earlier section were in textbooks reporting theoretical findings in this area, finding efficacy studies in the research literature was challenging. Following seven extensive literature searches using databases including: CINAHL, ERIC, ProQuest, PsychInfo, Cochrane Database, Health Source, PsychARTICLES, and Google Scholar, limited controlled and empirically based treatment research was located in peer-reviewed publications. This literature reported a paucity of empirically supported evidenced based practice (EBP) research for /r/ treatments. Most evidence was primarily at the level of expert opinion with a few case studies or single subject design studies; very limited small group studies were found. The studies found during the literature search are cited elsewhere within the text of this manuscript.

Theoretical Basis

Neurosensory motor system theoretical framework. Anita van der Merwe (1997) proposed a theoretical framework of sensorimotor control of speech by positing a four phase process for speech production including: linguistic-symbolic planning of a desired message; planning of consecutive motor movements; motor programming for the timing and spatio-temporal aspects of muscles for articulation, phonation, and respiration; and the motor execution of that plan. These four distinct phases account for the sensorimotor feedback loops and the communication that occurs in the brain at the neural level. Distinguishing four phases of sensorimotor function has helped the field of Speech-Language Pathology to better define and differentially diagnose disorders [i.e., linguistic planning deficits correlate to language disorders, motor planning deficits correlate to apraxia, motor programming deficits correlate to Parkinson’s disorder and some dysarthrias, and motor execution disorders correlate to articulation disorders]. Using this rubric, articulation disorders may result from a deficit of the core motor plan, which relies on sensory feedback (i.e., auditory feedback from production results, tactile and kinesthetic feedback in the sensorimotor memories, and refinements based on sensorimotor feedback loops) to make subtle corrections to the plan for accurate articulatory productions (van der Merwe, 1997).

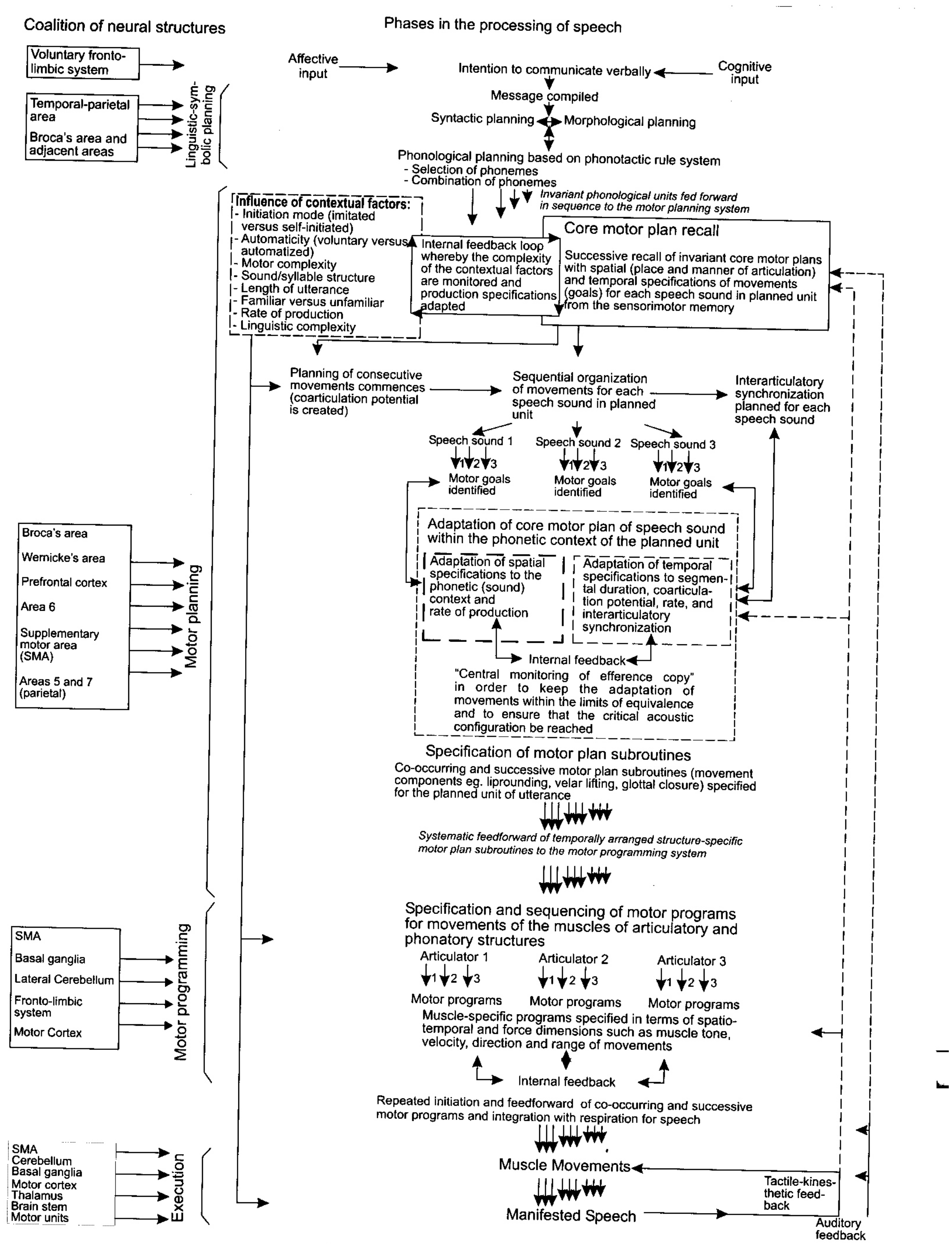


Figure 3. Four-level framework of sensorimotor control of speech production. Adapted from "A theoretical framework for the characterization of pathological speech sensorimotor control" by A. van der Merwe, 1997, In M. McNeil (Ed.), Clinical Management of Sensorimotor Speech Disorders. New York, NY: Thieme Medical Publishers, Inc., p.8.

This study examined how the use of a manual movement as visual and kinesthetic external feedback to supplement the client's internal feedback affects what van de Merwe called the motor program. This feedback helps to inform the force and spatio-temporal aspects of articulatory movements to improve the feedback loop and consequently, the articulatory production. This model is rooted in limb motor research and is based on the hypothesis that speech movements incorporate sensory feedback updates into the preplanned motor program (van de Merwe, 1997, as cited in Rusiewicz, 2010).

Principles of motor learning theory. When considering /r/ articulation errors as a motorically based deficit, one cannot neglect examining the theory of the principles of motor learning (PML). PML is rooted in the work of kinesiology. Therapies that incorporate PML have been demonstrated to be effective with various populations including: acquired apraxia of speech (Freed, Marshall, & Frazier, 1997; Rosenbek, 1985; Rosenbek et al., 1972; Square, Chumpelik, Morningstar, & Adams, 1986; Wambaugh, Kalinyak-Fliszar, West & Doyle, 1998, as cited in Edeal, & Gildersleeve-Neumann, 2011), CAS, and SSDs (Skelton, 2004, 2007; Skelton & Funk, 2004; Strand & Debertine, 2000; Strand et al., 2006, as cited in Edeal & Gildersleeve-Neumann, 2011).

PML includes the following intervention considerations: repeated practice of the target motor task (Rosenbek et al., 1973, as cited in Edeal & Gildersleeve-Neumann, 2011); blocked versus random practice; mass versus distributed practice; variability of practice (Ballard, 2001; Duffy, 2005; Schmidt & Wrisberg, 2004; Strand, 1995; Yorkson, Beukelman, Strand, & Bell, 1999, as cited in Edeal & Gildersleeve-Neumann, 2011); knowledge of results versus knowledge of performance (including extrinsic versus intrinsic feedback); and rate considerations (Schmidt & Wrisberg, 2004, as cited in Edeal & Gildersleeve-Neumann, 2011). Maas and colleagues (2008) summarized the six aspects of practice to consider during intervention: practice amount, practice distribution, practice variability, practice schedule, attentional focus, and target complexity. Considering the complexity of motor learning for speech, each of these six practice elements is explored below.

Blocked practice (treated by section in ordered phases) and constant practice (treated in the same way with the same target) is conducive to positive performance during a session but may not result in the generalization of learning into functional contexts. A large amount of practice, however, needs to be coupled with variability as conversational demands are unpredictable rather than rote. Variability is practicing different targets in varied contexts which effects skill transfer. In this study, /r/ will be practiced in isolation, CV form (vocalic /r/ contexts), CVC form, at the word level using a variety of contexts and positions of the target within the word (see Appendix A), and will be probed during a conversational speech sample every session.

Distributed practice timeframes are more beneficial than massed practice, by spreading learning out over time to solidify it. A random practice schedule, in which targets are randomly rotated into therapy sessions, is similar to stimuli of everyday interactions, accounting for the generalization of learned practice into functional settings. To promote generalization in this study, vocalic targets will be randomized.

Relevant literature has suggested that practice include complex targets as opposed to simple targets considering client experience in natural environments of target sounds in complex combinations and in all positions of words in continuous speech. According to current research, targets should become complex and functional as quickly as reasonable for clients to assist with generalization of targets. In this study, to address generalization, carry-over practice will be encouraged by instructing the participant to think about using manual mimicry during naturalized conversational speech settings.

Maas and colleagues (2008) discussed three feedback conditions that need to be considered: feedback type, feedback frequency, and feedback timing. Interestingly, knowledge of results (KR) appears more efficacious than knowledge of performance (KP). That is, it is more important for a client to receive feedback regarding whether a target production was accurate rather than how the production was generated. The premise is that clients need to develop internal understanding and representation of correct and incorrect productions without potentially confusing and/or perceptually incorrect feedback from clinicians. Lower amounts of delayed feedback appear more beneficial to clients in comparison to high levels of immediate feedback. These two findings emphasize the need for clients to actively construct an internal mechanism for analysis, storage, and retrieval of correct target productions and errors; additionally they highlight the need for clinician restraint and controlled feedback conditions. During implementation of this study, KR was given on a random schedule with a slight delay after productions, and the client will be asked to consider if the target was correct or incorrect prior to clinician providing KR.

Resource allocation theory. As discussed above, the clinician's feedback of target correctness is more beneficial to client retention than describing the mechanics or details of how that target was produced (i.e., placement of articulators, force of movement, voicing elements, etc.). Although, these details may be necessary during acquisition of a target, focusing the client on this level of detail may potentially overwhelm his/her system. The ineffectiveness of high levels of detailed feedback can be explained by McNeil, Odell, and Tseng's (1991) resource allocation theory (RAT), by individual production variability, or by the interference it causes with the client's ability to intrinsically assess and evaluate his/her target production.

Considering RAT, if the target speech production is difficult for a client to produce independently, it is assumed that the client requires large amounts of focus and resources to accomplish the task accurately. When detailed information about clinical observations is supplied, the client needs increased resources to process this information. However, processing this information may inappropriately draw from resources the client needs for production of the target. Individual production variability is well documented in typical speech production. A target description supplied by the clinician from his/her own perspective may be confusing for clients, challenging their ability to find an individual method of accurate target production. In this study, the focus is on the client simply understanding the representational correlation of the clinician's hand to the client's tongue. This will be the only verbal and foundational knowledge necessary for the client to understand the changes that need to be made based on the clinician's manual mimicry cue.

Self-cueing. The ability to self-monitor correctness and evaluate errors is vital during the process of acquiring a skill (see Maas et al., 2008). Clinician interference during this process undermines development of self-regulation and creates a situation in which clients may look to and rely upon clinicians instead of developing self-monitoring skills. This has obvious impacts on the functional generalization of skills when the clinician's presence is limited, and the goal of intervention is self-reliance. Concerning this study, the client's participation in determining correctness is expected following most productions and prior to clinician providing KR.

Early in treatment more support is necessary to achieve target productions, requiring more detailed, frequent, and motivational feedback. However, feedback type, frequency, and timing need to be adjusted once the skill begins to be acquired. The reasons for reduction of feedback include: over-reliance on clinician feedback and the reduction of client opportunity to process information, attend to sensory feedback, self-monitor, and categorize the learning into retrievable information for future performance. In this study, the first two treatment sessions will allot time for explaining representational elements of manual mimicry and providing limited verbal feedback accompanied by a visual, graphic, and pictorial representation of the manual mimicry cue hand position with a picture of the oral cavity to reinforce the correlation (see Appendix B).

Dynamic systems theory and entrained systems theory. The concept of self-cueing is consistent with the concept of entrainment from dynamic systems theory (DST).

DST postulates the value of coordination and integration of various pieces of information across systems and multisensory inputs from different contexts (Phillips-Silver, Aktipis, & Bryant, 2010). DST, in addition to PML, addresses target complexity, in which the complexities and dynamism of human beings is considered (Rvachew & Bernhardt, 2010). DST applied principles from the field of physics to a model of motor control. In dynamic systems theory, the speech, gesture, and language systems are intertwined. Therefore, the gestural theories based on this model attempt to explain the gestural mechanisms of speech as they relate to the entire language system. Because the focus of DST is on these fully integrated systems and language formulation, the basic tenets of gesture use for speech unrelated to language formulation are not clearly delineated. The reader is directed to Rusiewicz (2010) for a thorough review of gestural theories with relation to entrainment and a more thorough history of dynamic systems theory beyond the scope of this paper. The entrainment systems theory is the closest to generating a theoretical perspective for the type of gesture addressed herein.

Entrainment is defined as "spatiotemporal coordination resulting from rhythmic responsiveness to a perceived rhythmic signal…a coordinated rhythmic movement based on capacities for perception and production of rhythmic information, and the real-time transmission of this information between sensory and motor systems" (Phillips-Silver, Aktipis, & Bryant, 2010, p. 5). Iverson and Thelen (1999) proposed the theory of speech system entrainment grounded in dynamic systems theory by examining gesture use in infants and toddlers. The authors posit the entrainment or connection between hand and mouth activity at birth through the feeding mechanism (e.g., Babkin reflex) and deemed them "coupled oscillators" (p. 11). They postulated that "speech and gesture are temporally synchronous and are part of a unified system" (Rusiewicz, 2010). Iverson and Thelen (1999) discussed rhythmic movements becoming gestures, gradually entraining to the vocal system which, through practice, becomes "a tight synchrony of speech and gesture in common communicative intent" (p. 36). They address development of speech in relation to gesture as a coupled system and the motoric elements of gestures prior to linguistic overlay. An interesting tenet of this theory is that the gestural system is one of decreased complexity in developing children in comparison to the speech system. This basic tenet supports the use of a gesture, what is considered an earlier developing system in the dynamic system of human beings, to assist with the more complex speech system.

Clayton, Sager, & Will (2004) distinguish stimuli for entrainment by clarifying that an external element is not necessary because an individual can self-entrain by synchronizing two or more bodily systems. More research is necessary to determine the connection between limb and speech systems and the potential for entrainment; however, a clear connection is present between the two systems as evidenced by: verbal and non-verbal communication and speech and gesture coupling. Mayberry and colleagues (1998, as cited in Clayton, Sage, & Will, 2004) found that speech and gesture were so entwined that gestures ceased during stuttering moments. There is a lack of systematic investigations of gesture at the motoric level. This study proposes examining the functional effects of a specific gesture on a targeted speech sound; however, it does not propose to determine how those two systems are connected and interrelated. Further research into the area of motor speech and gesture entrainment is necessary and is beyond the scope of this project. The theoretical underpinnings considered in the development of the treatment used in this study include: speech sensorimotor control theory, PML, RAT, DST, and entrainment systems theories.

Significance

This project provides a fundamental contribution to /r/ articulation treatment by using a controlled research paradigm with limited confounding variables to parse out the targeted effect of the innovative manual mimicry gesture to improve articulatory accuracy of the /r/ phoneme. Ruscello's (1995) survey results expressly requested novel and improved therapy approaches for /r/ articulation disorders. This experiment will provide data on this novel approach to /r/ remediation. Secondly, the use of the EPG instrumental system solely as a dependent measure will contribute to literature for the use of this tool as an outcome measure. Subsequently, this study will indirectly contribute to gesture and motor speech entrainment literature perhaps providing insight for future areas for research into the exact mechanisms for entrainment of hand movements to oral articulators. Furthermore, the experiment elucidates future research directions as the first systematic investigation to specifically examine effectiveness of a hand gesture as a treatment for articulation disorders and specifically for /r/ errors. Finally, this analysis of clinically relevant and effective treatment strategies for /r/ errors contributes to the EBP literature available to practicing clinicians to assist with this challenging remediation and offers an enhanced view and additional tools to supplement traditional therapy techniques.

Research Methods

Purpose

The purpose of this experiment was to assess the influence of manual mimicry treatment on the accuracy of vocalic /r/ productions as measured by (i) listener perceptual judgments by a clinician, experts, and naïve listeners (ii) expert clinician transcriptions and (iii) EPG "gold standard" accuracy measures of tongue to palate contact during production of vocalic /r/ probes.

Experimental Design

This study was a single subject ABAB research design, meaning that baseline data was gathered, treatment administered, treatment was withdrawn during the maintenance phase, and finally, treatment was re-administered. Treatment was conducted during nine 60 minute sessions over a two month time span. Each session was conducted one or two times per week, as determined by the participant's schedule. Each session consisted of two distinct treatment periods (duration of 25 minutes each) with a short break (of 10 minutes) between sessions see Appendix C). The independent variable was the manual mimicry treatment. The dependent variables were listener perceptions (clinician, expert, and naïve) of accuracy, expert transcriptions, and EPG "gold standard" tongue to palate contact measures.

Treatment Participant. Participant enrollment began following approval of the research protocol by the Duquesne University Institutional Review Board. One treatment participant (female; 21 y.o.) was recruited for participation.

Recruitment. Recruitment of potential treatment participants was conducted with the Department of Speech-Language Pathology at Duquesne University and the larger Duquesne University community. Flyers (see Appendix C) were placed in the lobby of the Duquesne University Speech-Language-Hearing Clinic and provided to faculty of Duquesne University. Potential participants contacted the investigators via phone or email as indicated on the flyers. Current or past clients were contacted by the past and current clinical instructor for the Speech Production Clinic. Both males and females from any racial/ethnic background were invited to participate in this study.

Inclusion/exclusion criteria. Participant inclusion criteria was based on participant self-report and included: currently producing /r/ distortions as his/her primary speech sound disorder; a history of a speech sound disorder, including the inaccurate production of /r/; no more than three additional speech sound errors; over the age of twelve and under the age of 30; no interfering deficits of hearing, language, cognitive function, vision, reading, and/or oral motor skills; as well as no known or perceived concomitant medical diagnoses. Exclusion criteria included: existing orthodontia or oral prosthetics; organic articulation disorders; phonological disorders; language disorders; and atypical non-verbal IQ. Scores that indicated lower than anticipated performance (i.e., >1 SD below the mean) by a participant would not preclude participation in the study.

Participant. One individual, (female; 21 y.o.), a college freshman with English as her primary language, was recruited to participate in this project from the Speech Production Clinic affiliated with the Speech-Language-Hearing Clinic at Duquesne University. The study was described to the participant via phone as delineated in the script in Appendix E. A series of questions were also asked via this script to assess the eligibility of the individual for the study. At the time of initial phone call, the purpose and procedures of the study were explained and a brief series of questions were asked to verify eligibility for the study.

The participant presented with a speech sound disorder including inaccurate production of /r/, without demonstrating co-existing disorders (e.g., hearing, neuromotor, behavioral, cognition). Participant history included therapy services since early elementary school for speech sound errors /s, tʃ, ʃ/. Prior to enrollment in this study, the participant was receiving speech therapy twice per week for /r/ and learning support for reading comprehension. She was not enrolled in any other speech therapy program for the duration of this study.

Participant assessment and results. During the first two sessions formal and informal assessments took place to provide information on the individual's speech, language, vision, nonverbal intelligence, and oral motor skills. This was done to ensure that none of these variables would confound results. No results found during the assessment appeared to be capable of confounding treatment effect. Details of each assessment are described below.

Assessments. An audiometric hearing screening at one, two, and four kHz at 25 dB HL was conducted to ensure that hearing was within functional limits. Audiometric results determined that the participant had normal hearing. An informal vision screen using stimuli (with the same size and font as treatment probes) from the /r/ word list for the study (Ristuccia, 2006) was used to determine the participant's capability to participate in reading the stimuli (see Appendix F). The participant's vision was within functional limits. An informal assessment of the participant's voice was completed using a standard procedure; the participant's vocal quality, pitch, and intensity were within functional limits. The participant's rate of speech during conversation was noted to be rapid. An oral motor and mechanism examination was completed to examine the participant's structure for signs of organic articulation disorder and to examine function of the oral motor mechanism (oral mech exam adapted from Robbins and Klee's clinical assessment of oropharyngeal motor development in young children, 1987). Oral mechanism exams assess both the anatomy and physiology (i.e., form and function) of the structures used for speech sound production. Each feature is examined systematically to determine accurate function for speech. It is administered by asking the participant to perform simple oral motor tasks and demonstrating them as needed. Results of the oral motor exam revealed oral motor form and function to be within normal limits; however, subclinical differences included: slightly small lower mandible; high arched maxilla; teeth alignment in a U curve shape; tongue enlargement in the oral cavity; tongue rounded upon protrusion with an obstructed view of the velopharynx; lack of crispness in articulatory contacts during diadokokinesis; and mild hypernasal resonance. The participant's medical history was significant for tonsillectomy and childhood ear infections. The participant's language skills were formally assessed for consideration of any receptive language deficits that may confound treatment. A formal receptive vocabulary assessment was administered using the *Peabody Picture Vocabulary Test, Third Edition (PPVT-3, 1997).* The PPVT-3 is a norm-referenced formal assessment designed for individuals aged 2:6-90+ years that provides information regarding receptive (hearing) vocabulary attainment for Standard English that requires no speech or writing. It is administered by providing four pictures and asking the respondent to point to the word spoken. Results of the PPVT-3 revealed a raw score of 194 and a standard score of 95, resulting in a 37th percentile rank. A confidence interval of 95% was used to determine a range of standard scores from 86 to 105. The normal curve equivalent (NCE) was 43, the stanine was 4, the growth scale value (GSV) was 207, and the age equivalent score was 17:2 giving the participant an average score compared to normative data from peers in a similar age range. The participant's articulation abilities were formally assessed to determine current articulatory skill level and systematically assess all speech sounds for errors. The formal speech production and articulation assessment used was the *Goldman-Fristoe Test of Articulation, Second Edition* (GFTA-2, 2000). The GFTA-2 is a systematic means of assessing an individual's articulation of the consonant sounds of Standard American English. It provides a wide range of information by sampling both spontaneous and imitative sound production, including single words and structured conversational speech. It is administered by asking the respondent to name pictured items and to repeat stories with pictorial stimuli. Results of the GFTA-2 revealed a raw score of 11 and a standard score of 55, placing the participant in the <1 percentile rank with a test-age equivalent of 4:4. With a 95% confidence interval, the standard score range is 51 to 59. This is well below the expected performance for norms of age-matched peers. All errors noted during the assessment were with the /r/ sound except for two instances of lateralized /s/ during a cluster production of /sp/ and labialized /l/ during a cluster production of /pl/. The errors noted with /r/ included dehroticized /r/ during cluster production of /dr/ at the sentence level and /tr/ at the word level; labialized /r/ with productions at the sentence level in the medial and final word positions; at the word level in all positions and with the following clusters at the word level (/br/, /dr/, /kr/, /tr/). An error of note was a substitution error of r/w and kr/kw in the initial positions at the word level. The participant's non-verbal intelligence was examined to ensure that no confounding variables would interfere with comprehension for treatment tasks. The formal non-verbal cognitive assessment used was *the Test of Non-Verbal Intelligence, Fourth Edition (TONI-4, 2010).* The TONI-4 is a norm-referenced instrument that measures an individual's intelligence without using words by allowing responses to include pointing, nodding, gesture, and blinking during tasks of simple orally administered instructions. Results of the TONI-4 revealed a raw score of 36, an index score of 96, a rank in the 39th percentile, and an age equivalent of 14:6. Descriptively, this placed the participant in the average range.

Naïve listener recruitment. Naïve listeners provided "layperson" perception data on the accuracy of approximately 30% of the /r/ sounds produced by the participant. The naïve listeners were twenty-eight women between the ages of 18 and 30 (M = 21 y.o.) recruited within the Rangos School of Health Sciences Department of Speech-Language Pathology (SLP) at Duquesne University. They were recruited via flyers provided in SLP courses, in student mailboxes in the SLP Department, and posted in the student work areas within the department in Fisher Hall. Additionally, some participants were provided extra credit for their participation within a given course when approved by the instructor. Both a verbal announcement and written information regarding the investigation were given directly to the students of these courses. Participants responded to the flyer by contacting the PI or student Co-Investigator by phone or email, or by placing their name on a sign-up sheet posted in the SLP student resource workroom.

Naïve listeners were asked to provide their age, whether they had treated an individual with the primary objective of treating the /r/ sound, ratings of confidence in their dichotomous (correct/incorrect) judgments, level of participant naturalness, and qualitative feedback following the listening task. The data obtained from their dichotomous (i.e., correct/incorrect) judgments of the productions provided the opportunity to explore the relationship between expert transcriptions and EPG data, as well as offering perspective on the functional effect of the treatment to natural environments.

Expert listeners. Additionally, two expert listeners, faculty from the Speech-Language Pathology Department with a combined twenty-five years of experience phonetically transcribed approximately 10% of the data and performed dichotomous judgments on approximately 20% of the data.

Stimuli. Stimuli consisted of five vocalic /r/ index cards (AIR, ORE, ARE, IRE, EAR) and eighty /r/ words in varied contexts and positions within the word presented in size 85 Tahoma font. An illustration of the manual mimicry hand cue with an illustration of the oral cavity was provided for education during the first few sessions and it was also used as a visual reminder of the representation of the clinician's hand for the participant's oral cavity. See Appendices A, B, and F for stimuli examples and word lists.

Setting. All research was conducted at the Duquesne University Speech-Language-Hearing Clinic in a designated treatment room behind one way glass. The participant and clinician were inside the treatment room throughout the duration of each session, with the exception of the ten minute break.

Procedures. There are four distinct sets of procedures of this treatment study protocol (Appendix C). First, the participant completed the series of assessments of speech, language, nonverbal intelligence, and oral motor function delineated above. Second, the participant had a mold made of her palate by a local orthodontist and a customized palate created by Complete Speech for later electropalatography measures. Third, the participant completed the baseline and treatment procedures associated with the treatment of /r/ using manual gestures. Lastly, three types of listener rated randomized recordings.

Audio/video equipment. An audio and video recording using the Speech Production Clinic's *Landro* recording equipment was tested, set up, and utilized for every session. The lapel microphone distance was measured at maximum 50 cm from the participant's bottom lip. An additional wireless microphone headset (Shure PG30 microphone, PG1 Wireless transmitter, PG4 wireless receiver) was positioned three inches from the participant's lips and was used for high quality sound recordings for later naïve and expert ratings. Sampling rate was set at 48 kHz. All probes were recorded using the above equipment into Audacity recording software as wav. files with a project rate of 44100Hz.

Creation of customized palate for electropalatography. Electropalatography was employed as a dependent measure of tongue to palate contact during the production of /r/. Once informed consent was complete, the participant scheduled and attended an appointment with the designated local orthodontist, who was the identified provider for mold creation for both clinical and research endeavors in the Department of Speech-Language Pathology at Duquesne University. A mold was created at the orthodontist's office. The orthodontist took an impression or cast of the participant's upper teeth and roof of mouth (palate) that set within one minute. From this impression the orthodontist created a plaster stone model identical to the palate. From that cast, the Complete Speech technicians created an acrylic *SmartPalate* with electrodes imbedded in the palate (Figure 4).

Figure 4. Custom-made SmartPalate® for the Complete Speech Palatometer System. For more information visit:

http://www.youtube.com/watch?v=_AtiZxwTnpw&feature=player_embedded

The customized palate was then used for approximately 10-20 minutes during the baseline and treatment sessions to measure tongue contact to the palate during /r/ production (Figure 5). The Complete Speech EPG software program was run on a Dell laptop computer. The software and equipment (palates and wires) were tested and prepped, and a calibration was conducted for each session.

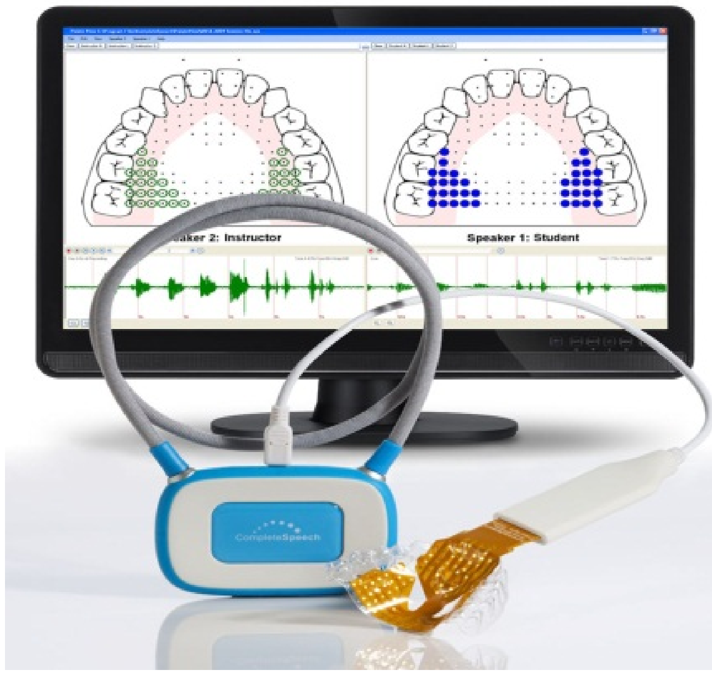


Figure 5. The Complete Speech EPG software system components. For more information visit: www.completespeech.com

The Complete Speech EPG software program visually demonstrates contacts made between the tongue and palate. This program also performs a calculation of the number of electrode sensors on the *SmartPalate* (placed on the palate in the mouth) that were touched by the tongue during the production of a sound. The EPG system consists of: the Palatometer software program (Palate View) (that was loaded onto the laptop computer); the Data Link microprocessor I/O device which connected that SmartPalate to the PC; and the custom-fit pseudopalate SmartPalate worn inside the mouth to sense mouth to palate contacts.

The audio recording from the EPG system (captured at 44 kHz) was stored in the EPG system as native wav. file format that was played back with any media player for future analysis in real-time (although slow and stop motion features are available). An integrated Omni-directional microphone internal to the Complete Speech EPG system is capable of basic sound acquisition at 20-60,000Hz; however, an external microphone was connected to the integrated mini-stereo jack for higher fidelity audio recordings. On screen EPG data views of the /r/ phoneme were exported as jpg. files of the EPG sensor data being recorded at 100 Hz or 100 times per second. The SmartPalate consists of 126 gold-plated contacts with 122 palate sensors, two lip closure sensors, and two gum contact sensors that are sampled 100 times a second. The Palate has an onboard microelectronic multiplex 124 unique channels over a shielded cable. The Data Link powers the EPG system when plugged into a standard computer using a USB cable. The SmartPalate is connected using the USB port at the top of the Data Link. Once the Palate View software was opened, the EPG system was functional and palate to tongue contact could be visualized. Electrodes on the screen were color-coded for visualization of correct targets (blue) and incorrect tongue-to-palate electrode activation (orange). EPG data from sessions was saved and exported into Excel for analysis.

Data collection procedures and session structure. There were a total of nine appointments yielding 18 discrete sessions for the treatment participant in which data was collected on the production of /r/. Baseline and assessment data were gathered over the first two appointments. The next two appointments provided manual gesture treatment, as described below, and conduct probes. The treatment was withdrawn for two appointments to assess generalization and maintenance of the trained skills during baseline measurements. The remaining three appointments resumed manual gesture treatment.

Independent and dependent variables.

Independent variable (IV). The IV of this study is the manual gesture treatment (Figure 6). The clinician conducted the manual mimicry hand cue by shaping the clinician's dominant hand within six inches of her cheek, directly paralleling tongue movement with the hand gesture. The clinician explained that the tips of her fingers symbolized the tongue tip and that her hand demonstrated the qualities, shape, configuration, and tension of the tongue at rest. The clinician then moved her hand synchronously with tongue movement during the /r/ phoneme into a cupped hand position, mimicking tension, placement, shape, orientation in mouth, trajectory and speed of movement (symbolizing the bunching and raising of the back of the tongue toward the hard palate) for /r/ production. The clinician's hand mimicked tongue placement for the vowel within the vowel quadrilateral (Appendix G) the movement of the tongue from that specific vocalic vowel into the /r/ sound for each of the five vocalic /r/ configurations was demonstrated. The clinician's manual mimicry gesture was explained, demonstrated, and performed simultaneously with any production of /r (clinician and/or participant's /r/ production).

Explanations of the representative aspects of the hand mimicking the tongue included: "my hand movement is imitating my tongue movement", "my hand represents the place of my tongue in my mouth", "my hand is moving with the timing of my tongue movement inside of my mouth during the /r/ sound", "for the /r/ sound, the back of my tongue rises up like this [use gesture to demonstrate]", "I feel the bunching and tension in my hand just like the bunching of my tongue".

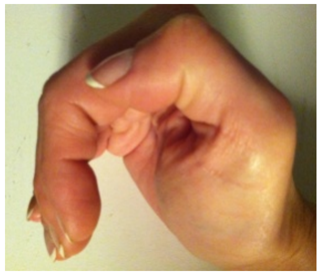


Figure 6. Example of a manual mimicry cue for /r/.

Dependent variable (DV). The dependent variables (DV) are (Figure 7):

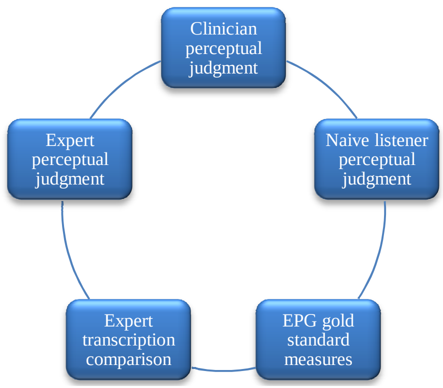


Figure 7. Dependent variables to determine /r/ production accuracy.

Defining dependent variables. Five measures were analyzed and compared; although, only three were specifically delineated as dependent variables (i.e., expert perceptual judgment and transcriptions, naïve listener judgments, and EPG measures). For the purpose of organizing the data, all five measures will be labeled as dependent variables (DV). The five DV are delineated below.

1. Clinician perceptual judgments (correct/incorrect) were made for following probes both on-line and during a second listen of a recording (audition only condition with </= two listens per target):
   a. 25 Baseline 1 word level productions
   b. 50 Baseline 1 vocalic /r/ productions

c. 125 Treatment 1 vocalic /r/ probes

d. 200 Baseline 2 vocalic /r/ probes

e. 30 Baseline 2 word level productions

f. 150 Treatment 2 vocalic /r/ probes

g. 16 Treatment 2 word level generalization productions

A total of 596 probes were analyzed twice by the clinician and an intra-rater reliability was calculated between those two listening conditions.

2. During the probes, the EPG palate was in place to capture "gold standard" measurements. The Complete Speech software program has a built-in "gold standard" electrode sensor contact pattern for each English language phoneme (i.e., the /r/ sound). The contact patterns during a participant's production of a sound were compared to this "gold standard" and a percent contact measurement was obtained from the software program.

3. Perceptual dichotic judgments (correct or incorrect) of productions of vocalic /r/ were made by three types of listeners (expert, naïve, and clinician). The three types of listeners reviewed randomized DVD recordings from the vocalic /r/ probes and recorded perceptual judgments on a data sheet provided (Appendix H). The expert listeners also transcribed a portion of these vocalic /r/ productions.

    a. The perceptual listening (judgment of correct or incorrect /r/ productions) of the clinician conducting treatment was recorded for all probes.

    b. The perceptual listening (judgment of correct or incorrect /r/ productions) of two expert clinicians was recorded for almost 20% of the recorded probes on provided data sheets, totaling 106 randomized data points of the total 596 tokens (18%).

   i. The on-line narrow transcription of incorrect /r/ productions by the clinician conducting treatment and the two expert clinician raters were recorded for almost 10% of the recorded probes on the data sheets provided, totaling 53 randomized data points of total 596 tokens (9%).

c. The perceptual listening (judgment of correct or incorrect /r/ productions) by twenty eight naïve listeners comprised of students in the Department of Speech-Language Pathology were completed for almost 30% of the recorded probes on the data sheets provided, totaling 184 randomized data points of the total 596 tokens (31%).

Procedures

Baseline and treatment sessions.

Initial baseline phase (A_1BAB). Baseline measures of /r/ productions were taken during sessions one through three and included:

1) A conversational speech sample of running speech was recorded for the first three to five minutes of each session. This conversational sample was used for informal assessment of voice and intelligibility at the level of conversation and provided a continuous baseline for improvement measurements.

2) Twenty-five contextual randomized /r/ sound probes were solicited during the first session in all word positions and levels of complexity (i.e., vocalic, consonantal, stressed, unstressed, blends/clusters, multisyllabic, in running speech, etc.) to gain a baseline to compare to future probes examining generalization of treatment (Appendix A). These probes were conducted without the EPG palate.
3) Twenty-five randomized vocalic /r/ sound probes (i.e., AIR, OR, AR, EAR, IRE) were solicited as a baseline measure. During the second session, the EPG SmartPalate was in place during twenty-five randomized vocalic /r/ sound probes (i.e., AIR, OR, AR, EAR, IRE) to obtain a measure of variance that may relate directly to the placement of the palate in the mouth.

First treatment phase (AB_1AB). The first set of treatment sessions using the manual mimicry gesture treatment began in session four and continued through session eight. Treatment sessions utilizing manual mimicry (Figure 6) included fifteen minutes of treatment followed by twenty-five randomized vocalic /r/ sound probes using the treatment strategies and cueing. Without the EPG palate in place, the microphone was positioned three inches in front of the participant's lips. The clinician explained the concepts of the manual mimicry hand cue delineated above when defining the IV.

The clinician slowly retracted her tongue synchronously with her hand movement when the phoneme /r/ was said in isolation, in five vocalic /r/ contexts (i.e., ER, AIR, AR, IRE, and OR), and at the word level. The clinician performed the hand movement simultaneously with the participant's /r/ production, as well as her own. The clinician increased speed to a natural and conversational rate at the level of VC (vocalic /r/) as mastery dictated (determined by 80% accuracy). The clinician slowly added levels of complexity to the participant's /r/ productions, when 80% accuracy was achieved at a given level. CVC level was beginning to be trained upon completion of the experiment.

Corrective placement visual cues, accompanied by limited simple verbal cues, were supplied as needed during the therapy training time and during the manual mimicry treatment probes (i.e., mimicry gesture hand movement and direct model). Unison productions with the clinician were performed as needed. The benchmark of 80% accuracy was determined based on the clinician's perceptual judgment data accompanied by the Complete Speech EPG "gold standard" electrode placement for the /r/ phoneme. When needed, the inter-judge rater expert 2 supplied correct versus incorrect judgments for comparison to avoid allowing the clinician to acclimate to the client's speech sounds and make false evaluations based on comfort with the participant's speech patterns and prediction of errors. /r/ productions were maximized during the treatment session to ensure massed practice for the benefit of new motor learning. Knowledge of results (i.e., told if the production was correct or incorrect) was given to the participant.

At the end of the training session, immediately following the manual mimicry treatment, the probes for treatment occurred. Prior to initiating probes, the participant placed the EPG SmartPalate in her mouth. A test of EPG function occurred briefly. Then, twenty-five probes of /r/ in the five vocalic /r/ contexts (i.e., AIR, OR, AR, EAR, IRE) were conducted (Appendix A). The clinician supplied the manual mimicry technique while the client verbalized the randomized list of twenty-five probes of /r/ in VC (vowel- consonant) vocalic /r/ contexts, given clinician coaching, as needed per discrete session. The participant was provided with a copy of all stimuli as home practice items.

Second baseline phase (ABA_2B). Sessions nine through twelve were conducted as a treatment withdraw or maintenance / generalization baseline measure to assess treatment generalization effects and establish treatment efficacy. During each treatment withdraw session, 100 randomized vocalic /r/ sound probes were elicited. In addition, fifteen contextual randomized /r/ sound probes were elicited in all word positions and levels of complexity. During this phase, probes were elicited without treatment cueing to assess generalization of treatment and treatment effects. The EPG was in place for all probes to serve as a DV outcome measure.

Second treatment phase (ABAB_2). Sessions thirteen through eighteen were treatment sessions identical to the first phase.

Results

Results of the dependent variables

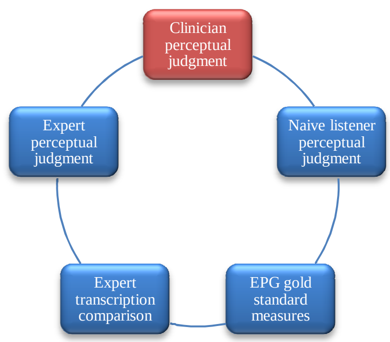


Figure 8. Aspects of dependent variables for results.

Clinician results. Twenty-five data probes of the vocalic /r/ productions were conducted by the clinician at the end of each of the eighteen sessions with the EPG in place. The clinician correct/incorrect vocalic /r/ judgments were averaged per condition to visually analyze mean variability (Figure 9) indicating the following mean, SD, and range (Table 1).

Table 1 *Phase mean, standard deviation, and range for clinician perceptual judgments of percent correct vocalic /r/ productions.*

	M	*SD*	Range
Baseline $_1$	36%	0.05	24 to 48
Treatment $_1$	63%	0.1	34 to 75
Baseline $_2$	48%	0.04	43 to 52
Treatment $_2$	75%	0.06	66 to 83

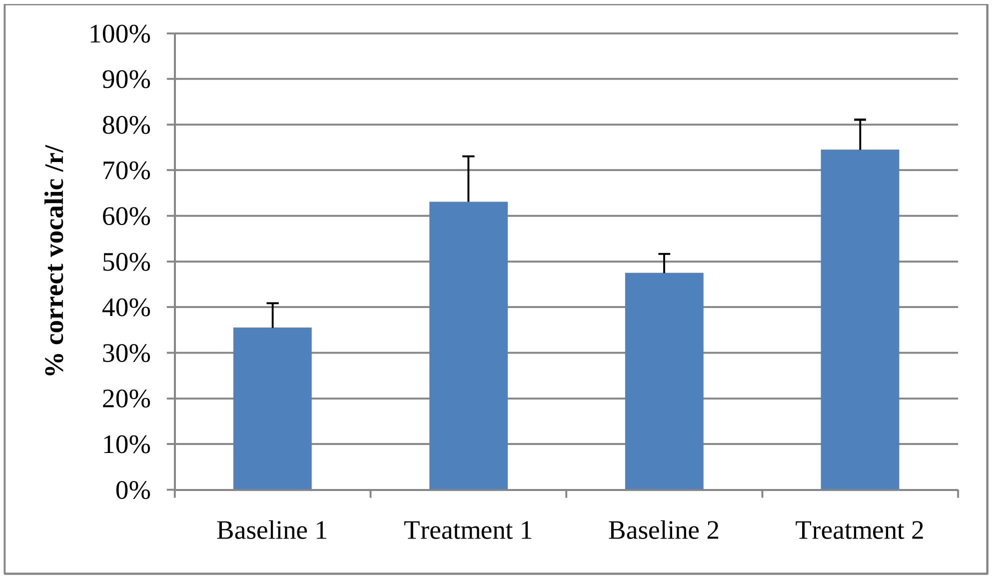


Figure 9. Clinician perceptual judgment means of % correct vocalic /r/ production averaged per condition.

Additional clinician findings. The clinician assessed the participant's overall conversational level comprehensibility and intelligibility. These data are presented in table 2.

Table 2. *Clinician perception of intelligibility and comprehensibility during conversational speech samples.*

Session	*Phase*	*Intelligibility rating*	*Comprehensibility rating*	*Notes*
On telephone	$Baseline_1$	75%	90%	
1	$Baseline_1$	85%	100%	Rapid speaking rate
2	$Baseline_1$	75%	100%	Very rapid speaking rate
5	$Treatment_1$	87%	100%	Facilitative context: produced the word "guard" clearly following the word "really"

6	$Treatment_1$	80%	100%	
7	$Treatment_1$	84%	100%	
9	$Baseline_2$	94%	100%	1st unquestionable /r/ production (not an approximation of /r/) "research"
11	$Baseline_2$	89%	97%	Very rapid speaking rate
13	$Treatment_2$	82%	98%	
15	$Treatment_2$	90%	100%	
17	$Treatment_2$	94%	100%	

The clinician assessed CVC (C+ vocalic /r/) accuracy during a probe in session 16, because the accuracy at the vocalic /r/ level was approaching 80% accuracy. CVC probes resulted in 68% accuracy /r/ production in that context.

The clinician assessed word level productions for /r/ sound accuracy (in all contexts, and positions of words) during session 12 as 33% accuracy; during session 17 as 43%; and during session 18 as 48% accuracy.

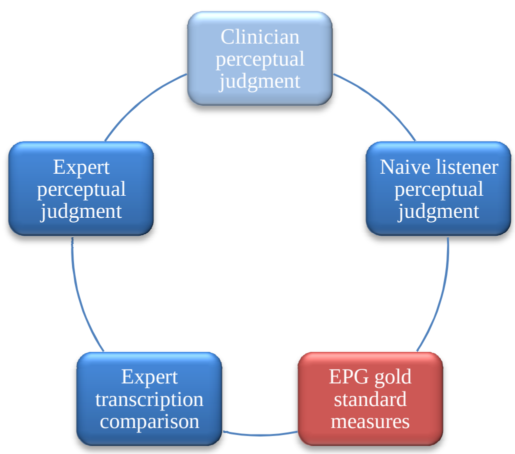


Figure 8. Framework for results from each dependent variable.

EPG DV results. The EPG results were examined to address the research question: Does manual cueing affect /r/ production as measured by electropalatography measures of percentage of accurate palate to tongue contacts?

EPG "Gold standard" for the /r/ sound as programmed by Complete Speech software calculated a percentage of correct tongue to palate (electrode contacts) for each /r/ sound made. EPG data was collected during the /r/ sound for all five vocalic /r/ productions per session. The mean for each condition was obtained beginning at the onset of the first treatment $_1$ session (*M* = 43.66%, *SD* = 0.009), baseline $_2$ (*M* = 46.25%, *SD* = 0.003), and treatment $_2$ (*M* = 45.68%, *SD* = 0.019) with an overall mean across all conditions (*M* = 45.27%, *SD* = 0.017). The EPG palate contact "gold standard" measurement of electrodes contacted by tongue to palate did not vary substantially from baseline to treatment (Figure 10).

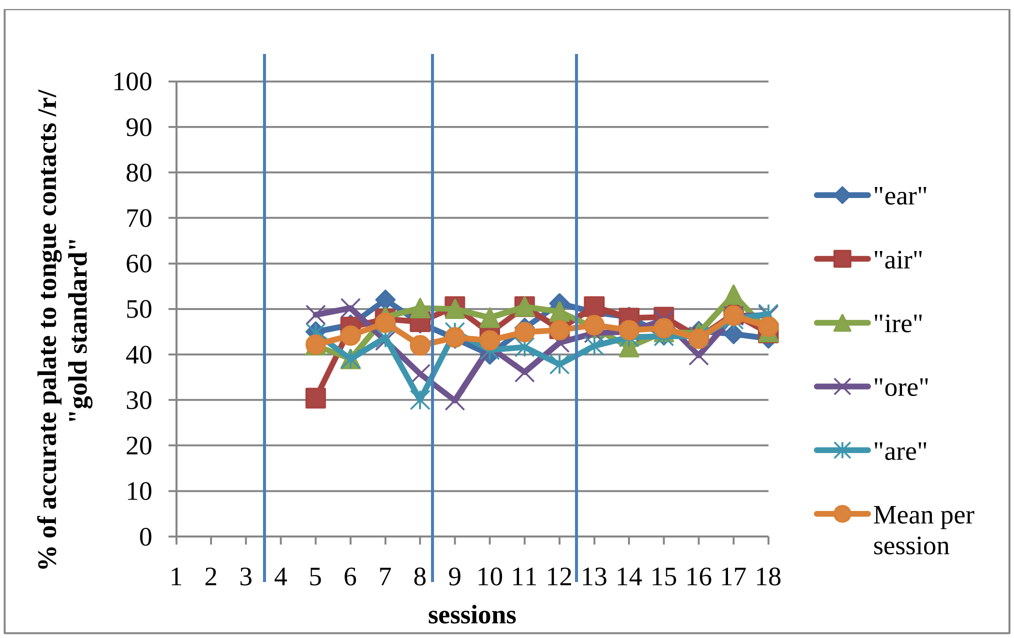


Figure 10. Complete Speech Electropalatography (EPG) "gold standard" for /r/ collected for 5 vocalic/r/ production probes per session.

Across all sessions in which the EPG "gold standard" measurements were obtained, the mean, standard deviation (M= 45.27%, SD= 0.02), and the range (42.49 to 48.62) of these measurements were minimal. Although, perceptual improvement of vocalic /r/ production occurred, it appeared from the EPG data that the tongue placement and positioning was not the primary reason for a perceptual change. Certain contexts for the VC vocalic productions appeared to be slightly more facilitative than others (i.e., AIR, EAR, and IRE appear to have slightly more appropriate tongue to palate contact; whereas, ORE and ARE remain lower in percent accuracy).

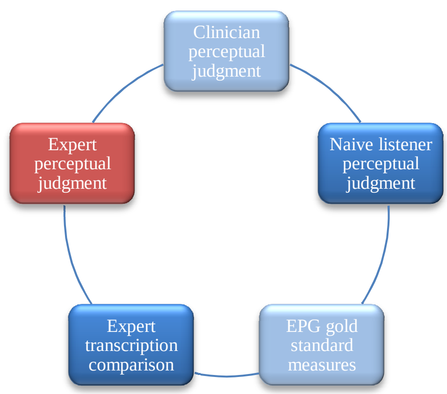


Figure 8. Framework for results from each dependent variable.

Expert results. The expert results were examined to address the research question: Does manual cueing affect /r/ production as measured by expert listeners' perceptual judgments of accuracy?

Expert inter-rater perceptual judgments. Inter-rater expert perceptual (correct/incorrect) judgments were made for a sample size of 106 randomized probes by listening to randomized DVD recordings of vocalic /r/ productions (no more than three times) in a quiet environment. Results of these judgments yielded per session accuracy by each expert listener as presented in Figure 11.

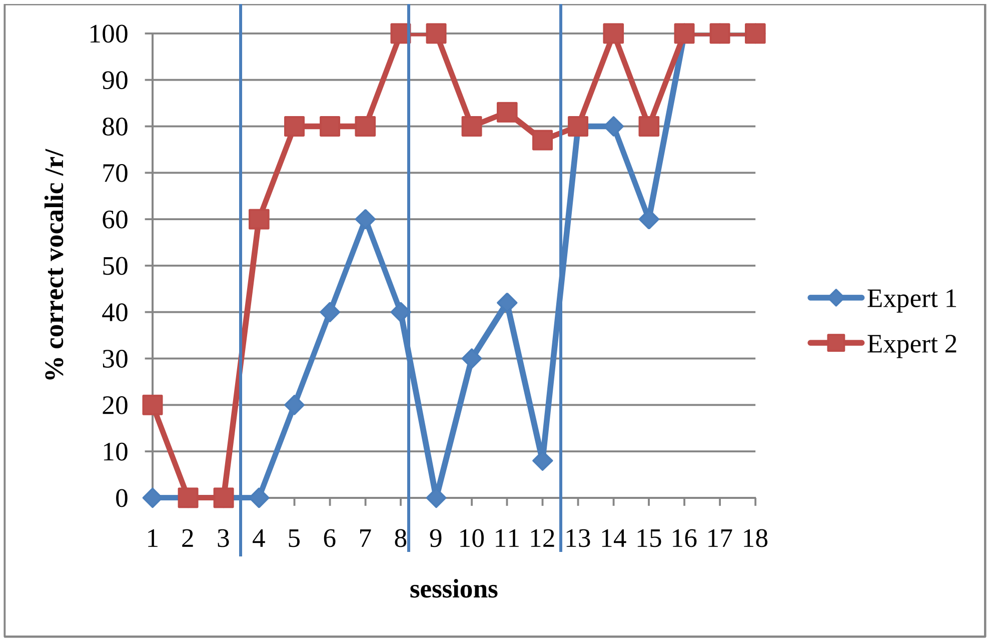


Figure 11. Two expert listeners' inter-rater perceptually judged vocalic /r/ productions dichotomously (correct/incorrect) yielding a percent correct per session**.**

The mean of the expert judgments revealed Baseline $_1$ (*M*=5%, *SD*=0.06364), Treatment $_1$ (*M* =56%, *SD* = 0.34), Baseline $_2$ (*M*=55%, *SD*=0.43), and Baseline $_2$ (*M*=90%, *SD*=0.05) (Figure 12).

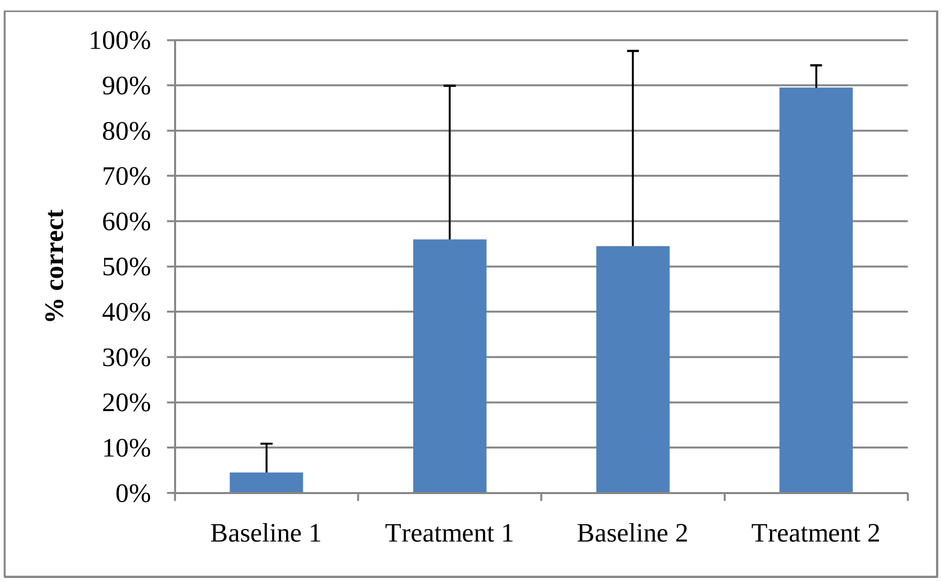


Figure 12. Expert perceptual judgments mean across conditions.

Comparing expert to clinician results. Results of the means per condition including both expert raters and the clinician indicate a perceptually stable baseline$_1$ (M= 3%, SD = 0.05), an increasing trend when the manual mimicry treatment$_1$ was initiated (M = 55%, SD = 0.24), a slight decrease in trend during the treatment withdraw$_2$ (M= 43%, SD= 0.36), and an increasing trend when the manual mimicry treatment$_2$ was re-instituting (M=89%, SD= 0.04), as indicated in Figure 13.

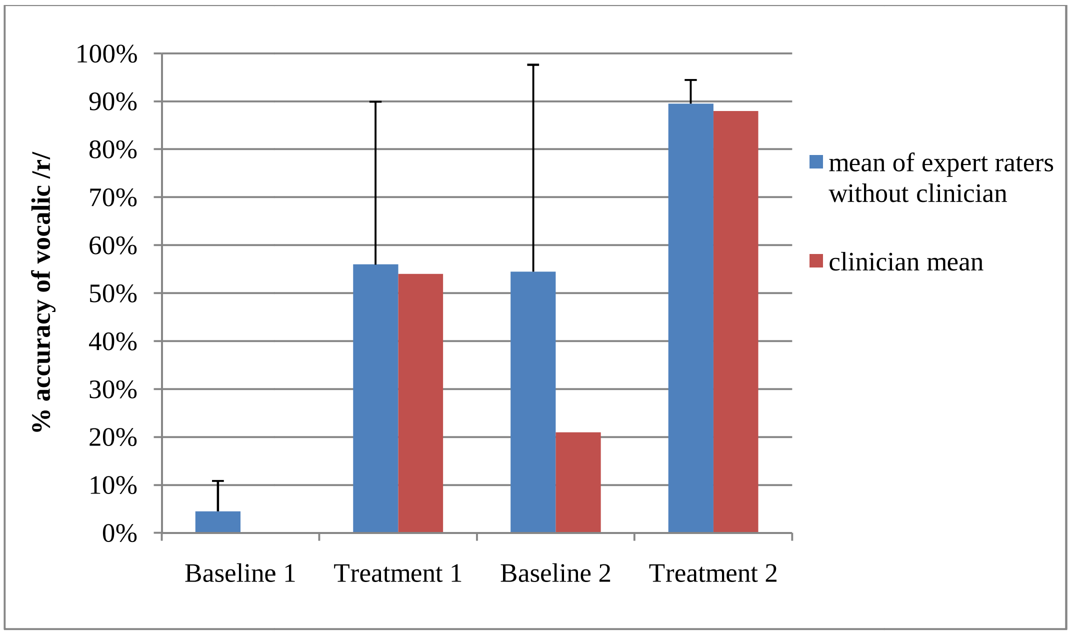


Figure 13. Mean percent accuracy of vocalic /r/ production dichotomous judgment of expert raters compared to the mean of the clinician per condition.

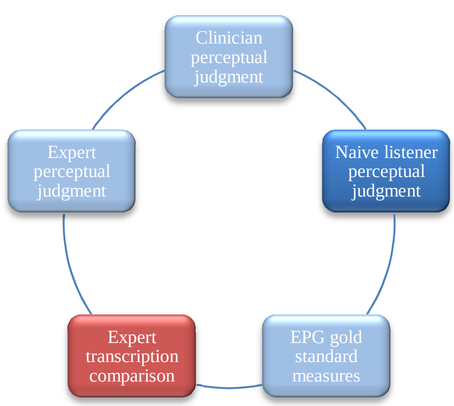


Figure 8. Framework for results from each dependent variable.

Transcription results. Broad transcription of fifty-three data points was conducted by two expert listeners (and one graduate student clinician), while listening to blinded randomized DVD recordings of vocalic /r/ probes (no more than three times) in a quiet environment.

The most prominent error type recorded was derhotization of /r/ in the final position of the vocalic CV context (10 occurrences of 53, 19% as judged by both expert listeners in agreement; and 26 occurrences of 53 as judged by either expert listener, 49%). Other error types of note included: one sound substitution in the initial position (w/r), lengthening of sounds (both vowels and /r/), addition of sounds (/ə/, /j/, and /h/), and retraction of /ɔ/. The characteristic qualities of the participant's /r/ errors were derhoticizing, lengthening, and addition of sounds; whereas, no omissions, and only one sound substitution were noted. See Appendix I for transcription details per listener.

Comparing expert listener transcriptions to clinician transcriptions. The clinician transcriptions were conducted online during the study for every speech sound production and later during a second listen of the 184 randomized productions, where narrow transcription error types were noted. One difference between expert and clinician judgment was the presence of sound substitutions vs. labialization of /r/. One expert perceived substitutions of /w/ for /r/; whereas, the clinician coded those productions as labialized /r/ instead which Shriberg and Kent (2003) have termed "nearly functionally equivalent". They suggest that this functional equivalence be used to set reliability strictness criterion. The clinician observed that baseline productions did not perceptually contain the wavering or lengthening qualities perceived in the other three phases. The lengthening of sounds (18 of 50; 36%) and wavering voice (11 of 50; 22%) were the most common perceptual features during treatment$_1$. Of note during that phase was that approximations appeared (slight dehrotization of /r/) (3 of 50; 6%), as did vowel distortions (2 of 50; 4%). The most salient features of the baseline$_2$ productions were approximations (9 of 70; 13%), vocal wavering (5 of 70; 7%), lengthening of sounds (4 of 70; 6%), sound additions (2 of 70; 3%), one instance of decreased volume (1 of 70; 1%), and one instance of decreased length (1 of 70; 1%). The salient features of treatment$_2$ were lengthening (24 of 48; 50%), the first instances of hard onset of sounds (18 of 48; 38%), increased volume (15 of 48; 31%), wavering (7 of 48; 15%), and distorted vowels (2 of 48; 4%).

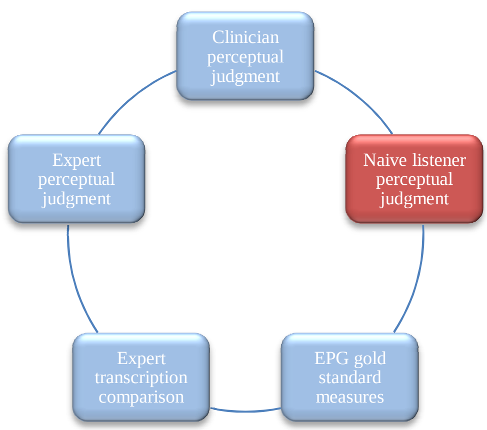


Figure 8. Framework for results from each dependent variable.

Naïve listener perceptual judgment results. The naïve listener results were examined to address the research question: Does manual cueing affect /r/ production as measured by naïve listeners' judgments of accuracy?

At the end of the participant's ABAB phases, randomized recordings and data collection sheets were generated for listener judges. Following informed consent procedures as approved by the Duquesne University Institutional Review Board, twenty-eight naïve listeners judged 184 randomized vocalic /r/ probes while listening to a DVD with each stimulus repeated twice. Up to five naïve listener participants at a time were seated in a quiet room free of distractions. They were provided a document with a list of syllables and words listed to guide their listening and judgments. They were instructed to listen carefully to the syllables and words that were played and to check the appropriate box to indicate that they were either completely correct or were incorrect in any way (Appendix H). The items were presented in a randomized order from the baseline and treatment sessions. An investigator played each item a total of two times for the listeners. This entire procedure, including the consent process took no longer than forty-five minutes per session.

The twenty-eight naïve listeners provided information about whether they had previous experience treating /r/ articulation errors specifically. The graphical breakdown of means separating those naïve listeners with experience (17 individuals; *N*= 61%) and those without experience (11 individuals; *N*= 39%) is presented in Figure 14.

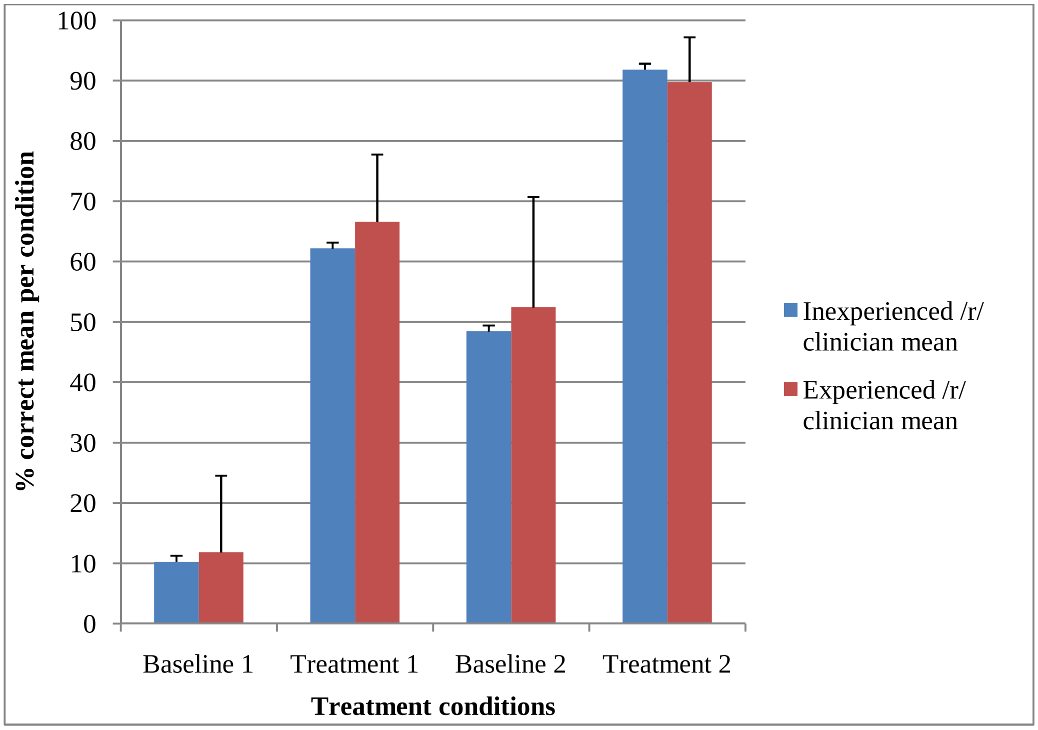


Figure 14. Mean naïve listener perceptual judgments of randomized vocalic /r/ productions per condition and segmented by listener experience treating /r/.

Naïve listeners were also asked to rate their confidence in the judgments they had made, resulting in a mean confidence level (*M* =79.74%, *SD* =4.12, *range* = 75% - 85%). Then, they were asked to rate the naturalness of the participant's speech production, resulting in a mean naïve listener naturalness judgment, with most listeners rating the productions as "fair" in naturalness and a few using the rating "fair to good".

The means across all naïve listeners per condition were baseline $_1$ (*M* = 11%, *SD* = 13.07), treatment$_1$ (*M* = 65%, *SD* =13.56), baseline $_2$ (*M* = 51%, *SD* = 19.38), and treatment $_2$ (*M* = 91%, *SD* = 6.83) (Figure 15).

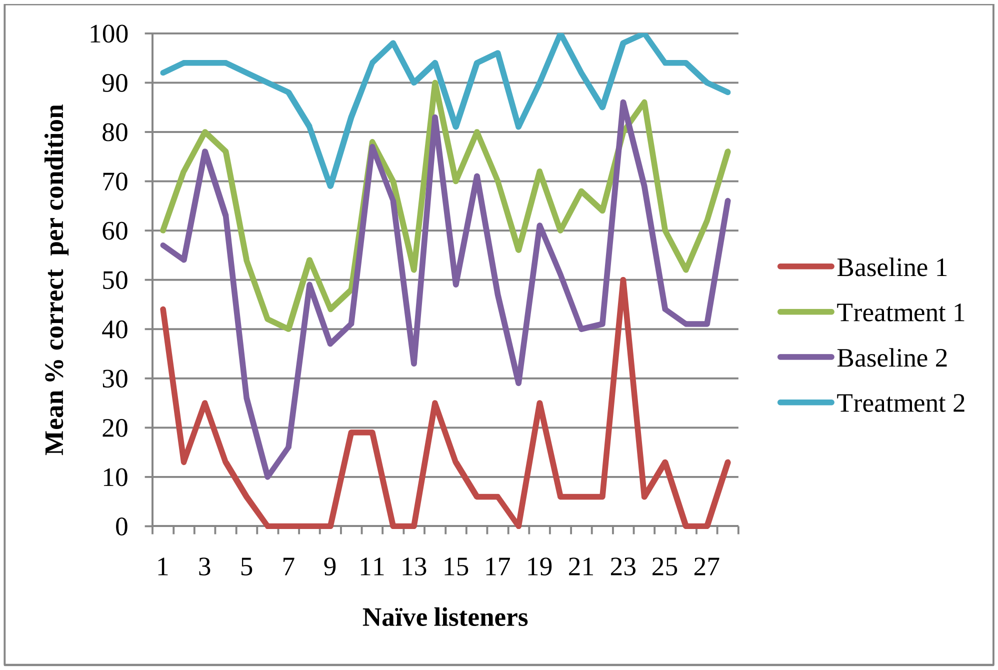


Figure 15. Mean of all naïve listeners' perceptual judgments of randomized vocalic /r/ productions per condition.

These data indicate high levels of agreement among the twenty-eight naïve listeners in the treatment $_2$ phase suggesting a treatment effect. Further visual analysis of these and all three types of listener perceptual results are presented below using visual analysis.

Rationale for visual analysis of single subject design data

Utilizing the visual analysis gold standard for single subject research design as specified in the What Works Clearinghouse in Kratochwill and colleagues. (2010), the dependent variables are represented graphically and analyzed below. Visual analysis of data is used to examine if evidence exists to link the presentation of the independent variable and the dependent outcome measures and to examine the strength of that relationship (Herson & Barlow, 1976; Kazdin, 1982; Kennedy, 2005; Kratochwill, 1978; Kratochwill & Levin, 1992; McReynolds & Kearns, 1983; Richards, Taylor, Ramasamy & Richards, 1999; Tawney & Gast, 1984; White & Haring, 1980, as cited in Kratochwill et al., 2010). A causal relationship is supported when data across the phases show three demonstrations of effect at three separate points in time. Visual analysis rules involve four steps and six variables. Step 1: confirm a predictable baseline pattern; Step 2: examine within phase patterns for consistency and predictability; Step 3: compare data from adjacent phases for evidence of the independent variable manipulation effect (associated with a predictable pattern of change in the dependent variable); and Step 4: integrate information across phases to determine if three demonstrations of effect occurred at three separate points in time to indicate a functional or "causal" relationship (Kratochwill et al., 2010).

Graphs of the data from this ABAB design were used to assess (1) level, (2) trend, and (3) variability within a similar data series and (4) immediacy of effect, (5) degree of overlap, and (6) consistency of data series between conditions. "Level" is the mean of the data within a phase. "Trend" is the slope of the best fitting straight line within a phase. "Variability" is the range or SD around the best fitting line. Within-phase data examination is conducted to describe observed patterns and predict expected performance, given no change to the independent variable (Kratochwill et al., 2010).

Across phase data examination is conducted to document a causal relationship inferring that the outcome variable was directly affected by the manipulation of the independent variable. "Immediacy of effect" is the change in level between the last three data points in one condition to the first three data points in the next condition, and rapidity of effect substantiates the inference that a manipulation of the independent variable had a direct effect on outcome measures. “Degree of overlap” is the portion of data in one condition that overlaps with the previous condition and the smaller the overlap the more indicative it is of a treatment effect. Percentage of overlapping data (POD) is calculated by determining the range of data in the baseline_1 phase, counting the number of data points in the treatment_1 phase, counting the number of data points of the treatment_1 phase that fall within the range of the first condition, and dividing the number of data points that fall within the range of the first condition by the total number of treatment_1 data points and multiplying this number by 100 (Gast, 2010, p. 214). “Consistency of data in similar phases” is examining phases of similar conditions with one another (i.e., all baseline phases) for consistency of data patterns within similar phases with the assumption that the greater the consistency, the more likely a causal treatment effect occurred (Kratochwill et al., 2010).

Data reduction using visual analysis across dependent variables.

Visual analysis of clinician judgments. Step 1: Demonstrate predictable and stable baseline pattern.

Concern is sufficiently demonstrated in baseline behavior by presenting a stable occurrence of percent correct vocalic /r/ productions below the anticipated ability of typical young adult /r/ productions, 31%, 42%, and 34% respectively (Figure 16). A behavior in need of remediation is clearly defined and consistent in level and variability, allowing comparison with treatment phase conditions.

Step 2: Analyze within phase elements.

Level. The mean percent correct vocalic /r/ production score within the $baseline_1$ phase is 36%, within the $treatment_1$ phase is 63%, within the $baseline_2$ phase is 48%, and within the $treatment_2$ phase is 75%. The data demonstrate consistency and predictability of pattern (Figure 16).

Trend. The trend of the two treatment phases as indicated by visualizing the slope of the best-fitting straight line demonstrates progression toward improvement of productions; whereas, the baseline remains relatively stable (Figure 17).

Variability. The range of the standard deviation around the best-fitting straight line $baseline_1$ phase is 42.57%; 32.43% (*M*= 36%, *SD*=0.57), within the $treatment_1$ phase is 75.10%; 47.90% (*M*= 63%, *SD*=0.1), within the $baseline_2$ phase is 52.04%; 42.96% (*M*= 48%, *SD*=0.04), and within the $treatment_2$ phase is 83.07%; 65.93% (*M*= 75%, *SD*=0.07) (Figure 18).

Step 3: Compare adjacent phases.

Immediacy of effect. The observed effects are immediate in comparisons across all phases ($Baseline_1$ to $treatment_1$, $treatment_1$ to $baseline_2$, and $baseline_2$ to $treatment_2$) analyzing the level, trend, and variability of the final three data points in one phase compared to the initial three data points for the adjacent phase. To more clearly demonstrate if any immediacy of effect (of instituting treatment) has occurred, these three points in time across each phase were visually compared; three different shapes have been superimposed onto those three points in the graphs below. The data in the ovals, rectangles, and triangles demonstrate immediacy of effect (Figure 19).

Degree of overlap. There is no overlap between data points from baseline $_1$ to treatment$_1$ or from baseline $_2$ to treatment$_2$. There are two overlapping data points (22.22%; sessions 9 and 10) between treatment$_1$ to baseline$_2$ (Figure 20).

Consistency across similar phases. The data patterns of similar phases indicate a consistent response of behavior under similar conditions as demonstrated in the linked ovals (Figure 21).

Visual analysis of mean clinician perceptual judgments within phases.

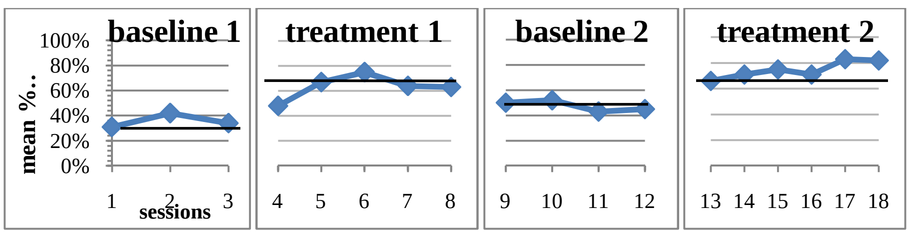


Figure 16. Visual analysis of level within phase for clinician judgment of vocalic /r/ production.

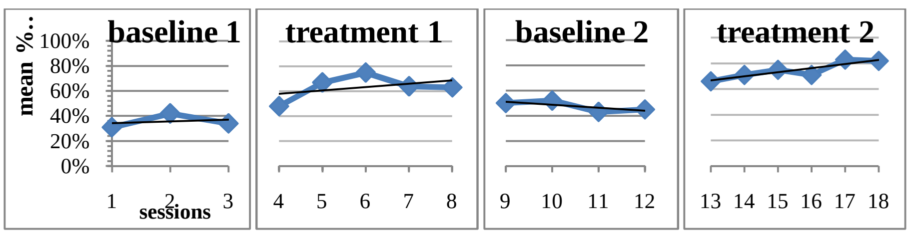


Figure 17. Visual analysis of trend within phase for clinician judgment of vocalic /r/ production.

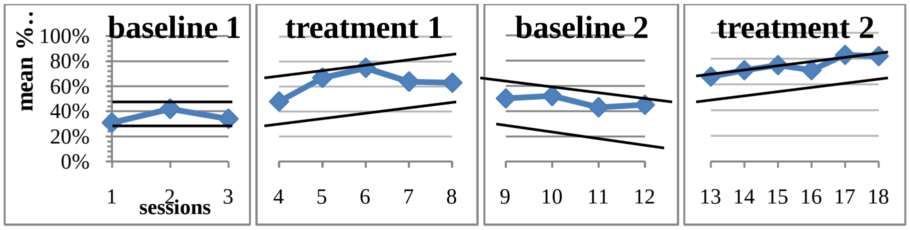


Figure 18. Visual analysis of variability within phase for clinician judgment of vocalic /r/ production.

Visual analysis of mean clinician perceptual judgments across phases.

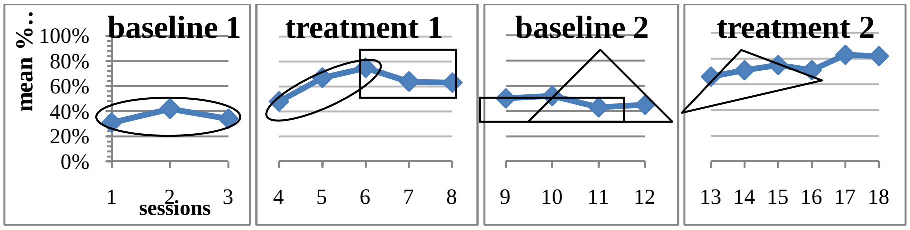


Figure 19. Visual analysis of immediacy of effect across phases for clinician judgment of vocalic /r/ production.

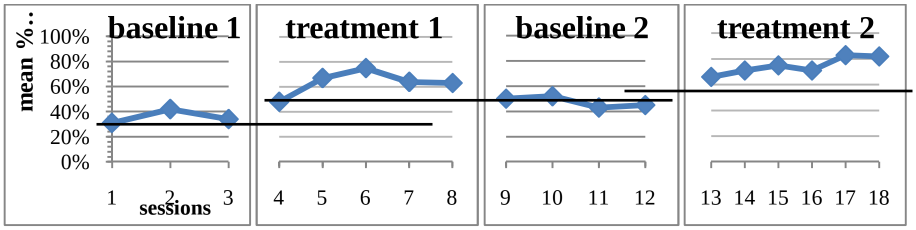


Figure 20. Visual analysis of degree of overlap across phases for clinician judgment of vocalic /r/ production.

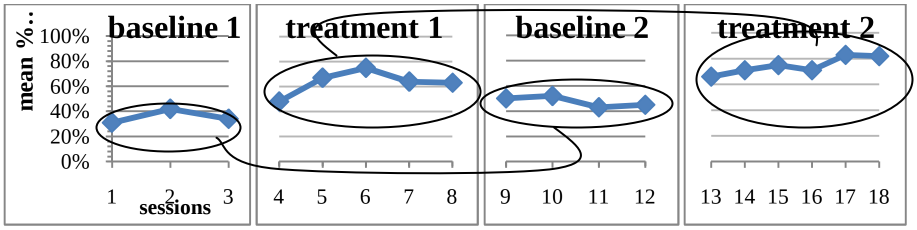


Figure 21. Visual analysis of consistency across similar phases for clinician judgment of vocalic /r/ production.

Visual analysis of expert judgments. Step 1: Demonstrate predictable and stable baseline pattern.

Baseline behavior presents a stable occurrence of percent correct vocalic /r/ productions well below the anticipated ability of typical young adult /r/ productions, 10%, 0%, and 0% respectively (Figure 22). A behavior in need of remediation is clearly consistent in level and variability, allowing comparison with treatment phase conditions.

Step 2: Analyze within phase elements.

Level. The mean percent correct vocalic /r/ production score within the baseline$_1$ phase is 3%, within the treatment$_1$ phase is 56%, within the baseline$_2$ phase is 53%, and within the treatment$_2$ phase is 90%. The data demonstrate consistency and predictability of pattern; although, the baseline$_2$ was not as low as anticipated for a withdraw condition (Figure 22).

Trend. The trend of the two treatment phases as indicated by visualizing the slope of the best-fitting straight line demonstrates progression toward improvement of productions; whereas, the baseline remains relatively stable (although more stable in the baseline$_1$ phase then in the baseline$_2$ phase (Figure 23).

Variability. The range of the standard deviation around the best-fitting straight line baseline$_1$ phase is 10%; 0% (*M*= 3%, *SD*=0.58), within the treatment$_1$ phase is

70.17%; 29.83% (*M*= 56%, *SD*=0.17), within the $baseline_2$ phase is 63.08%; 42.92% (*M*= 53%, *SD*=0.08), and within the $treatment_2$ phase is 100%; 69.87% (*M*= 90%, *SD*=0.13) (Figure 24).

Step 3: Compare adjacent phases.

Immediacy of effect. The observed effects are immediate in comparisons across all phases ($Baseline_1$ to $treatment_1$, $treatment_1$ to $baseline_2$, and $baseline_2$ to $treatment_2$) analyzing the level, trend, and variability of the final three data points in one phase compared to the initial three data points for the adjacent phase. To more clearly demonstrate if any immediacy of effect (of instituting treatment) has occurred, these three points in time across each phase were visually compared; three different shapes have been superimposed onto those three points in the graphs below. The data in the ovals, rectangles, and triangles demonstrate immediacy of effect between the oval and the triangles (Figure 25).

Degree of overlap. There is no overlap between data points from baseline $_1$ to $treatment_1$ or from baseline $_2$ to $treatment_2$. There are four overlapping data points (44.44%; sessions 9, 10, 11, and 12) between $treatment_1$ to $baseline_2$ (Figure 26).

Consistency across similar phases. The data patterns of similar phases indicate a consistent response of behavior under similar conditions (baseline $_1$ and baseline $_2$) as demonstrated in the linked ovals (Figure 27).

Visual analysis of mean expert perceptual judgments within phases.

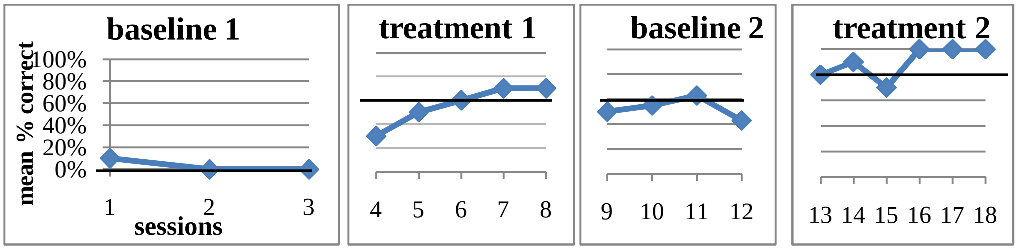


Figure 22. Visual analysis of level within phases for expert judgment of vocalic /r/ production.

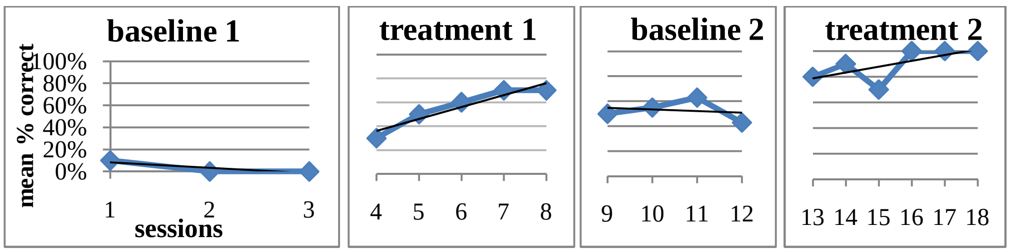


Figure 23. Visual analysis of trend within phases for expert judgment of vocalic /r/ production.

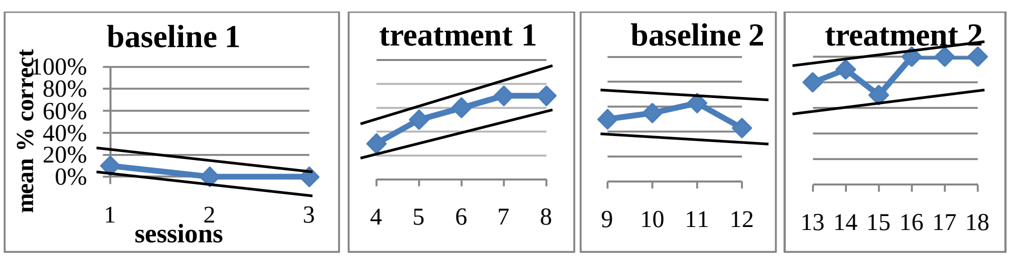


Figure 24. Visual analysis of variability within phases for expert judgment of vocalic /r/ production.

Visual analysis of mean expert perceptual judgments across phases.

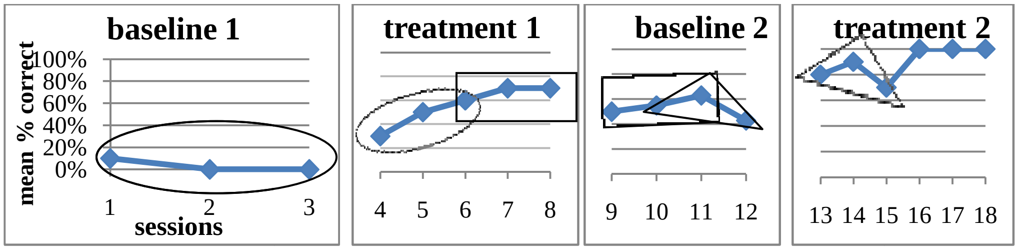


Figure 25. Visual analysis of immediacy of effect across phases for expert judgment of vocalic /r/ production.

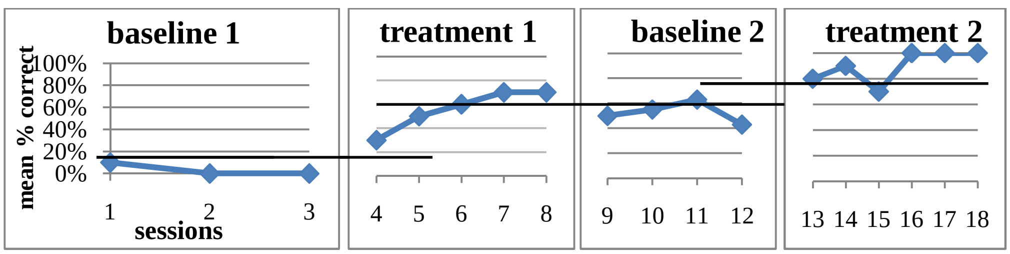


Figure 26. Visual analysis of degree of overlap between phases for expert judgment of vocalic /r/ production.

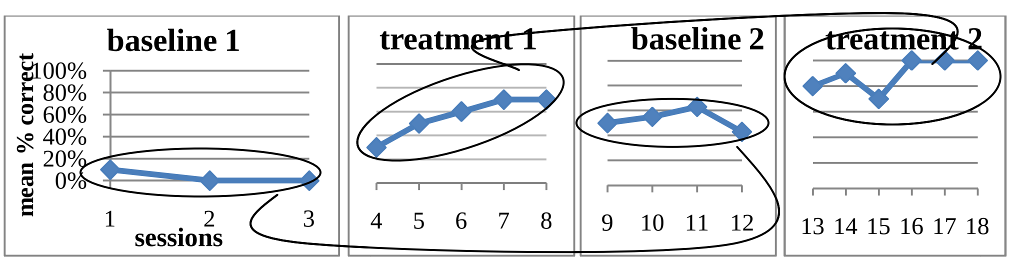


Figure 27. Visual analysis of consistency across similar phases for expert judgment of vocalic /r/ production.

Visual analysis of naïve listener judgments. Step 1: Demonstrate predictable and stable baseline pattern.

Baseline behavior presents a stable occurrence of percent correct vocalic /r/ productions below the anticipated ability of typical young adult /r/ productions, 13%, 11%, and 10% respectively (Figure 28). A behavior in need of remediation is clearly consistent in level and variability, allowing comparison with treatment phase conditions.

Step 2: Analyze within phase elements.

Level. The mean percent correct vocalic /r/ production score within the baseline_1 phase is 11%, within the treatment_1 phase is 65%, within the baseline_2 phase is 53%, and within the treatment_2 phase is 90%. The data demonstrate consistency and predictability of pattern (Figure 28).

Trend. The trend of the two treatment phases as indicated by visualizing the slope of the best-fitting straight line demonstrates progression toward improvement of productions; whereas, the baseline remains relatively stable (Figure 29).

Variability. The range of the standard deviation around the best-fitting straight line baseline_1 phase is 11.02%; 8.98% (*M*= 11%, *SD*=0.02), within the treatment_1 phase is 81.02%; 54.89% (*M*= 65%, *SD*=0.11), within the baseline_2 phase is 64.09%; 41.89% (*M*= 53%, *SD*=0.09), and within the treatment_2 phase is 98.07%; 79.93% (*M*= 90%, *SD*=0.07) (Figure 30).

Step 3: Compare adjacent phases.

Immediacy of effect. The observed effects are immediate in comparisons across all phases (Baseline_1 to treatment_1, treatment_1 to baseline_2, and baseline_2 to treatment_2) analyzing the level, trend, and variability of the final three data points in one phase compared to the initial three data points for the adjacent phase. To more clearly demonstrate if any immediacy of effect (of instituting treatment) has occurred, these three points in time across each phase were visually compared; three different shapes have been superimposed onto those three points in the graphs below. The data in the ovals, rectangles, and triangles demonstrate immediacy of effect between the oval and the triangles. The difference in position of the ovals in this graph demonstrates especially immediacy of effect between the baseline$_1$ and treatment$_1$ phases (Figure 31).

Degree of overlap. There is no overlap between data points from baseline $_1$ to treatment$_1$ or from baseline $_2$ to treatment$_2$. There are two overlapping data points (22.22%; sessions 9 and 11) between treatment$_1$ to baseline$_2$ (Figure 32).

Consistency across similar phases. The data patterns of similar phases indicate a consistent response of behavior under similar conditions (baseline $_1$ and baseline $_2$) as demonstrated in the linked ovals (Figure 33).

Visual analysis of mean naïve listener perceptual judgments within phases.

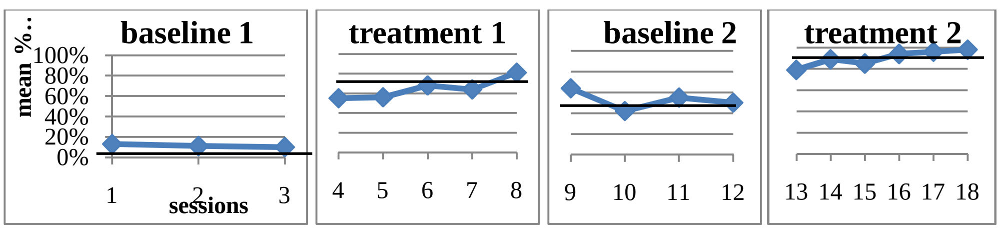


Figure 28. Visual analysis of level within phases for naïve listener judgment of vocalic /r/ production.

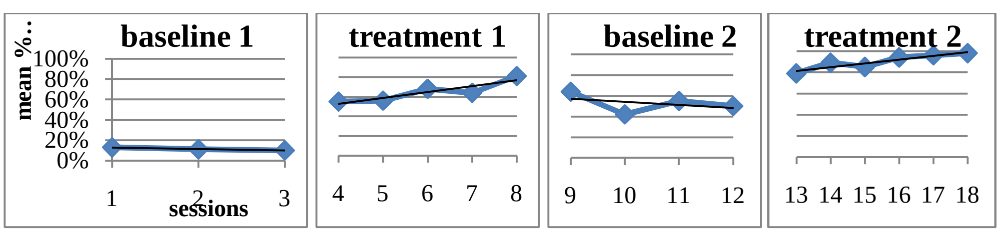


Figure 29. Visual analysis of trend within phases for naïve listener judgment of vocalic /r/ production.

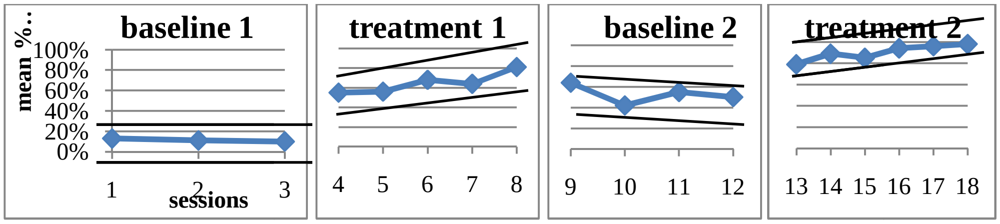


Figure 30. Visual analysis of variability within phases for naïve listener for clinician judgment of vocalic /r/ production.

Visual analysis of mean naïve listener perceptual judgments across phases.

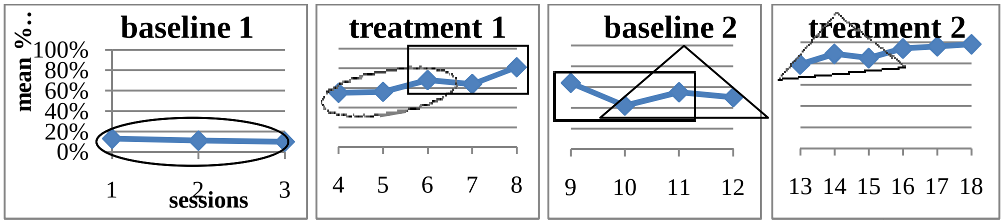


Figure 31. Visual analysis of immediacy of effect across phases for naïve listener judgment of vocalic /r/ production.

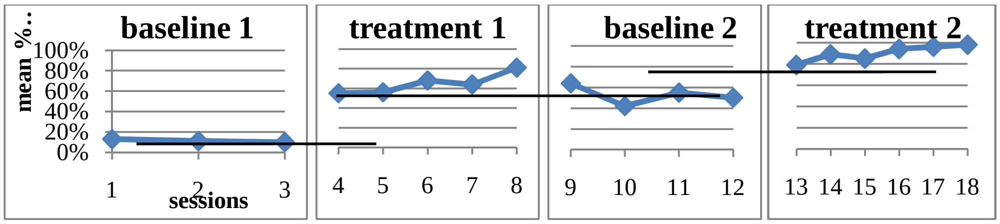


Figure 32. Visual analysis of degree of overlap between phases for naïve listener judgment of vocalic /r/ production.

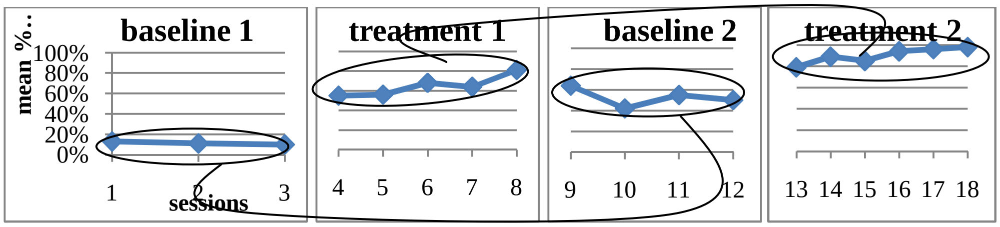


Figure 33. Visual analysis of consistency across similar phases for naïve listener judgment of vocalic /r/ production.

Reliability

Clinician Intra-rater reliability. The clinician made on-line judgments of correct and incorrect productions. Intra-rater accuracy was determined by re-listening to the probes using headphones (no more than two times each) and judging the probes a second time. The results and mean of these intra-rater accuracy judgments are graphically represented in Figure 34.

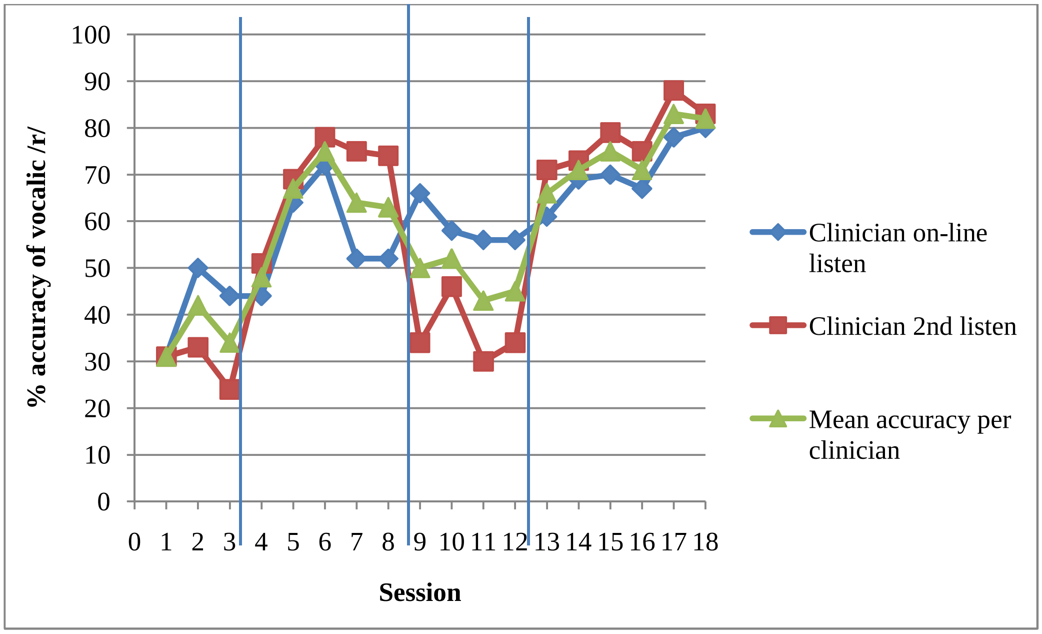


Figure 34. Clinician intra-rater perceptual correct/incorrect judgments of vocalic /r/ productions averaged per session by % accuracy.

The intra-rater reliability was calculated by comparing the clinician's on-line perceptual correct/incorrect judgments to the clinician's second time listening using headphones and recordings from the session to determine stability of the clinician judgments. Clinician intra-rater reliability was 65% for /r/ vocalic percent accuracy and 69% for /r/ word level percent accuracy.

Expert inter-rater reliability. Inter-rater reliability was calculated between the two expert judges across sessions by examining unit by unit agreement for perceptual judgments of 106 vocalic /r/ productions (Figure 35). Percent agreement was calculated using unit by unit agreement index= A/A+D*100 (whereas A= number of units agreed upon and D= number of units disagreed upon). Inter-judge reliability was lowest during the baseline$_2$ phase (session 9) with high levels of agreement in both the baseline$_1$ and treatment$_2$ phases.

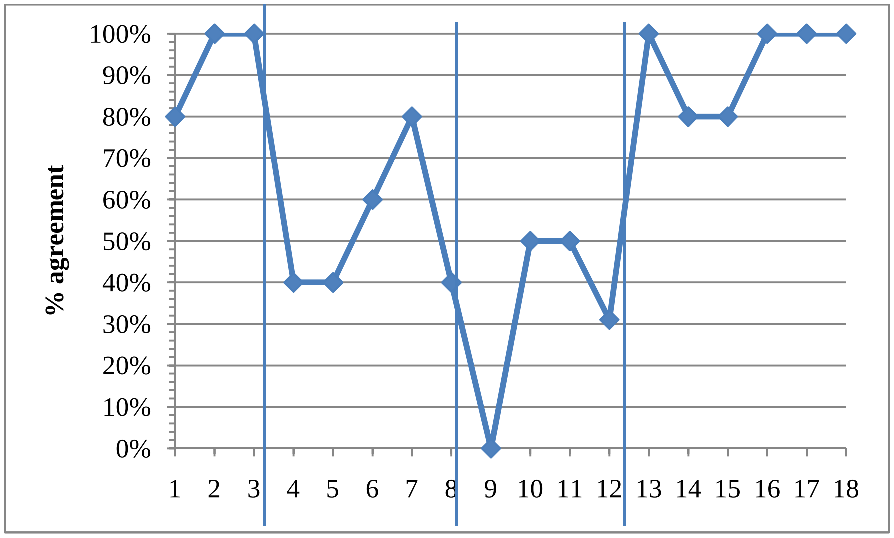


Figure 35. Inter-judge (expert 1 and expert 2) reliability measure of percent agreement for perceptual judgments of vocalic /r/ productions.

Expert inter-rater phonetic transcription reliability. The 53 narrow phonetic transcriptions for vocalic /r/ targets for both expert raters were entered into the Logical International Phonetics Programs (LIPP) software using an International Phonetic Alphabet (IPA) keyboard. The software TiteLipp setting was used for inter-rater reliability comparisons by calculating agreement between rater transcriptions using the KRELIAB.LAX rule. The inter-rater total of broad transcription consonant agreement was 0.98; broad transcription vowel agreement was 0.93. The overall broad transcription agreement between the two expert judges was 0.90. The total agreement of broad transcriptions between expert 1 and the clinician was 0.92, and the total agreement of broad transcriptions between expert 2 and the clinician was 0.96. The expert inter-rater agreement for narrow phonetic transcription of 53 /r/ productions was 40% agreement for when the diacritic markings were examined for error type production patterns (Appendix I). The intra- and inter- rating reliability were within acceptable ranges as indicated by previous research in this area (McSweeny & Shriberg, 1995; Shriberg & Kent, 2003).

Summary of results

Quantitative results. As previously discussed, there were three main questions addressed by this study. First, does manual cueing affect /r/ production as measured by electropalatography measures of percentage of accurate palate to tongue contacts? The EPG measures remained relatively stable throughout all conditions and demonstrated minimal movement upon initiation of the treatment. Therefore, according to this measurement, /r/ production was not affected by the use of manual mimicry during treatment.

Second, does manual cueing affect /r/ production as measured by expert listeners' judgments of accuracy of vocalic /r/ syllable contexts? Expert listener perception of accuracy was directly affected by manipulation of the IV for this participant. As expected, the effect of treatment was demonstrated by the increase in accuracy upon initiation of treatment as indicated by the judgments made by the raters (i.e., clinician, experts, and naïve listeners). It should be noted that the two expert listeners exhibited low agreement in the withdraw condition relative to the other three conditions.

Third, does manual cueing affect /r/ production as measured by naïve listeners' judgments of accuracy of /r/ in vocalic syllable contexts? The naive listeners, who represent the general population, provide a measure of the functional effect of this manipulation on this participant's performance. The perceptual accuracy by this group of listeners appeared to demonstrate a strong treatment effect, given the stable baselines and predictable increases upon manipulation of the IV.

A treatment effect was found based on the judgments made by the three types of listeners; the perceptual percent correct judgments significantly increased across all three listener types when the manipulation variable, namely in the $treatment_2$ condition was introduced. The conclusion of a positive treatment effect is strengthened by the consistent response and changes in rating between the three listener types. No treatment effect was found using EPG as a dependent measure.

All three listener types listened to 106 identical productions from the randomized samples. When these three listener means were compared for these data, the results corroborate a treatment effect based on the difference between $baseline_1$ and treatment $_2$ and again between baseline $_2$ and treatment $_2$. Results are presented in Table 2 and Figure 36.

Table 2

Mean percent correct vocalic /r/ for the three listener types across condition, average, and standard deviation of all listeners

	Clinician	Expert listeners	Naïve listeners	*M*	*SD*
	n= 1	*n* = 2	*n* = 28	*n* = 31	
Baseline 1	0%	5%	14%	6%	7.09
Treatment 1	56%	56%	66%	59%	5.77
Baseline 2	29%	54%	54%	45%	14.15
Treatment 2	86%	90%	87%	88%	2.08

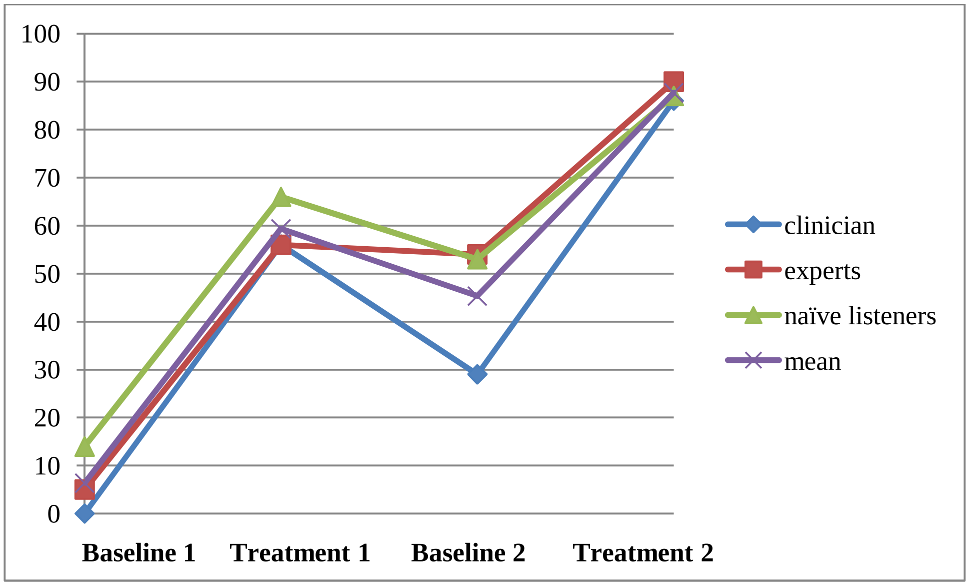


Figure 36. Mean percent correct vocalic /r/ for each listener type across conditions and averaged with all three listener types.

Qualitative findings. A post-hoc qualitative analysis was completed to examine emerging patterns from the three types of listeners (clinician, expert, and naïve listeners). These data were accumulated as clinician on-line notes during sessions, notations on forms by expert listeners, and open-ended questions asked to the naïve listener groups.

Clinician qualitative findings. *Baseline$_1$ phase.* The clinician only noted one instance of the participant's vocal wavering sound production during the baseline data collection process.

Treatment$_1$ phase. Beginning in session 4, a portion of probes were noted to have a tremulous nature and wavering quality toward the end of the vocalic productions (four of 30 productions) and other productions were executed correctly but at a low volume (two of 30).

Starting in session 7, the participant spontaneously began self-correcting productions and judging them as accurate or inaccurate on her own with requests for feedback on questionable productions. The clinician facilitated this self-monitoring behavior by randomly requesting participant judgment of accuracy prior to providing feedback. The participant offered the feedback that in her perception the "ARE" sound was the most difficult for her to produce accurately. It was noted that the participant began practicing in unison when the clinician was providing a direct model as well as performing self-rehearsal. The participant stated "seeing you and mimicking you" [helps make it easier to produce correct /r/]. During session 7, the participant also began to speed up productions of the vocalic /r/, which resulted in instances of vowel distortion. As of session 8, the participant requested "Can I fix that one?" and internal rehearsal continued. The wavering quality noted earlier perceptually lengthened and at times sounded as though the participant was adding a /ə/ sound, increasing the difficulty of accuracy judgments. Specific (and repetitive) verbal cues were given at this point (i.e., "tighter in the back" or "higher in the back") in addition to demonstrations of movement, placement, and tension of the tongue using manual mimicry cues. The clinician observed that the participant spontaneously generated the manual mimicry hand movement during the treatment condition, similar to the treatment effect that was demonstrated inadvertently in the Hall study (1992). However, the participant's hand movement occurred in various positions in space around her body and was not conducted directly next to her cheek or exactly mirrored to the clinician's hand position.

Baseline$_2$ phase. In session 9, the clinician was not providing the manual mimicry gesture during the baseline 2 phase. However, in that session, the participant began to intermittently spontaneously performing the hand gesture to self-cue productions. She also began using a faster rate as compared to the cueing condition (vowel distortions occurred). Increased approximations of accurate /r/ sounds were noted. In session 10, approximations were noted with slight to very slight errors and vowel substitutions were noted adjacent to correct /r/ productions. Self-rehearsal, addition of the /ə/ sound, and vocal wavering continued. The participant independently reported a generalization finding, explaining that her friend had noticed a change in being "able to understand" [her] "better when [she] 'does that hand thing'." In sessions 11-12, the participant's spontaneous hand gestures increased (again no clinician model provided during treatment withdraw/generalization phase). Vowel substitutions and distortions were noted, even with accurate /r/ productions. Approximations continued and rate changes were noted ranging from lengthening to rushing. During lengthening, /r/ sounds would at times become dehroticized toward the beginning of the sound or toward the end of the /r/ sound, not remaining consistently correct throughout one vocalic production.

Treatment $_2$ phase. During session 13, increased self-reliance was encouraged by the fading of cues (i.e., "find it" to indicate that the client's oral articulators should follow the hand movement to obtain proper mimicry, resulting in a correct /r/ sound. Manual mimicry cues were used to communicate movement from the vowel to the /r/ (namely, tongue height and tension in the production of "ORE") with limited verbal cues ("it starts in the back"). During session 14, the most remarkable finding was an increase in intensity. The participant spontaneously increased volume and tongue tension during vocalic /r/ productions. In session 15, the clinician slowly began to increase the speed of the model and probes while decreasing verbal cueing. Participant self-rehearsal, increased volume, and increased tongue tension continued to be noted. In session 16, an increase in volume continued. In session 17, increased initiation of utterances and utterance length in conversation was noted. Rapid rate of speech, increased volume, increased ability to self-correct, and more precise repetitions following a direct clinician model were also noted. Vowel distortions prior to correct /r/ productions continued to occur intermittently.

Expert listener qualitative findings. During transcription, expert listener 1 remarked on a "tremor and breathy" vocal quality.

Naïve listeners' qualitative findings. Following each naïve listener session the investigator asked them open-ended questions about the challenges they encountered while listening and rating this participant's /r/ speech sound productions. The naïve listeners were also invited to give general thoughts about the experience.

Some themes emerged from these discussions including: difficulty judging productions, specific techniques used to make judgments, differences in productions at varied levels (vocalic syllabic vs. word level), naturalness of participant's productions, and salient characteristics of the /r/ sounds heard (i.e., elongation, exaggeration, emphasis, increased volume, tremor, vowel distortions and differences between specific vocalic sounds (i.e., "IRE" vs. "ORE"). See Appendix J for detailed naïve listener observations.

Generally, the naïve listeners found the task challenging due to various factors of 1) the amount of variation in the production of the phoneme, 2) the difficulty of judging vocalic sounds with little context as opposed to making judgments at the word level, and 3) having to create a method of keeping an accurate /r/ model and keep it mentally on-line to have a basis of comparison for correct/incorrect judgments. The general consensus among naïve listeners included that "IRE" was the most difficult to judge, that

“ORE” seemed to be the most difficult for the participant to produce, that “EAR” was the most consistently accurate participant production, that lengthier productions or productions with vowel distortions were more difficult to judge, that louder productions (increased volume and effort) appeared to be more accurate, and that the productions overall were unnatural sounding.

Discussion

Summary of results

In this single-subject study, listener judgments indicated that manual mimicry cues had a positive treatment effect on intractable /r/ in a relatively short period of therapy. Within nine hours (18 sessions), vocalic /r/ productions had reached the level of mastery (over 80%) and CVC /r/ words were close (68%). The participant was beginning to be trained at the word level upon completion of the study.

Challenge of judging /r/. This study elucidated some of the challenges in treating /r/ errors. The first challenge is the variability in /r/ productions resulting in a spectrum of acceptable approximations and a possible gradation of the effect of therapy. Without repeated exposure to target productions, thorough analyses, and diligent monitoring of subtle changes, clinicians may have difficulty making consistent judgments. This challenge was best evidenced by the clinician's difficulty judging sound production targets (namely in sessions 8 and 9) due to increased approximations of /r/ sounds, requiring constant clinician adjustment and learning of the variations within this individual.

The second challenge pertains to determining an accurate and reliable method of measuring outcome success for the production of the /r/ sound. It was originally hypothesized that EPG would serve as a reliable outcome measure, secondary to its reported success as a treatment tool. However, in this study of the production of vocalic /r/ productions, EPG was not an adequate outcome measure. This participant only presented minimal change in EPG contact patterns (although some vocalic contexts were nominally more facilitative than others [i.e., AIR, EAR, and IRE appeared to have slightly more appropriate tongue to palate contact; whereas, ORE and ARE remained slightly lower in percent accuracy]). Yet, perceptually across three sets of listeners, a change in perceptual saliency of her /r/ productions occurred.

More success may have been possible by measuring /r/ in isolation or utilizing EPG for visualizing unspecified measurements in conjunction with the participant to see dynamic changes and gather qualitative data. However, as an outcome measure using the percent contact metric it was not as useful as anticipated. Another suggested methodological change would be to use frozen EPG screen shots showing placement of contacts, as seen in studies by Gibbons (1999), in which she evaluated "undifferentiated lingual gestures" by viewing screen by screen EPG pictures.

One hypothesis arose out of the lack of change in the placement of the tongue to the palate as measured by EPG. As stated earlier, although perceptual improvement of vocalic /r/ production occurred, it appeared from the EPG data that the tongue placement and positioning could not account for and therefore was not the primary reason for a perceptual change. It appeared in this particular participant, as in Gibbon's (1995) findings that the tongue was grossly in place and making the contacts with the palate appropriately. What appeared to be lacking was not placement but tension and tongue shape. It is hypothesized that her tongue went from lax and "undifferentiated" to tense and configured into a shape that allowed appropriate formant production in the vocal tract. This is an area for future research and also a caution for clinicians not to overemphasize tongue placement in /r/ remediation.

The third challenge was variability in the perceptions of the listeners and the reported difficulty that listeners had judging accuracy, given the numerous confounding factors during the /r/ productions. One factor in this difficulty was decreased production consistency and predictability as the participant learned and practiced new motor execution patterns; hence, greater variance in production to production became evident, requiring the clinician (and listeners, in general) to acclimate to the fluctuating targets. It is possible that judgments at the word level would be easier to discern than at the syllabic level due to context, as suspected by the naïve listeners. Potential evidence with respect to this possibility is provided in the clinician judgments of untreated levels (e.g., word level); these items have higher agreement ratings during the probes. Clinician intra-rater reliability was slightly higher for word level percent accuracy (69%) than for vocalic percent accuracy (65%). It was initially anticipated that the expert listeners would demonstrate the least variance between one another, but that they would have greater discrepancy from the other raters. However, this was not the case. All three listener types generally agreed on overall accuracy changes across most phases. The exception was the withdraw phase which had confounding results; the percent accuracy did not decrease as much as predicted, especially per expert listener 2. This baseline_2 was not as low as anticipated in a withdraw condition for level or trend. One potential explanation for this is that one expert judge had extensive transcription experience for research purposes (well over 300 listens) and is an instructor of Phonetics, which may account for the some of the differences in perception and transcription. Differences in the listener perceptions during the treatment withdraw phase ($baseline_2$) are hypothesized to be most likely due to: varied levels of experience; varied scoring for approximations; or potentially the participant's propensity toward independently self-cueing using the manual mimicry cue inconsistently throughout the withdraw phase which may have affected certain productions more than other productions.

The final challenge also relates to the making judgments about the accuracy of /r/ productions. If dichotomous judgments were difficult to align perfectly, it stands to reason that agreement of transcriptions would also be difficult to obtain. The qualitative feedback was useful for understanding specific aspects of this difficulty more extensively. The participant exhibited distortion of vowel sounds, prolongation, "undifferentiated lingual gestures" (Gibbon, 1995; Goozee, Murdoch, Ozanne, Cheng, Hill, & Gibbon, 2007), volume differences, and difference at the level of the larynx (e.g., hard onset and vocal tremor). All of these factors (and more) contributed to the difficulty reaching phonetic transcription agreements. It is well-known in the field of speech-language pathology that transcriptions vary based on clinician experience, training, comfort, and knowledge and that agreement is limited in any phonetic transcription task (Shriberg & Kent, 2003). It was expected, however, that clinicians within the same facility would have high degrees of agreement. Whereas, the agreement attained in this study was within an acceptable range, it was slightly lower than predicted. The attainment of agreement appears to have been confounded further by the overall difficulty of making judgments of the phoneme, /r/.

It is interesting to note that naïve listener experience with treating /r/ appears to correlate with increased variability and standard deviation around the mean, which was also seen with the expert listeners. Perhaps the fine tuning of trained ears to /r/ variance allows experienced listeners to perceive more subtle approximations of a correct /r/ sound. Having a "model" /r/ target production prior to listener judgments may have assisted them and decreased the burden on them to create internal models (per naïve listener qualitative feedback). This supports the use of naïve listeners as the gold standard in articulation and phonological disorders used to predict generalizability. Having untrained listeners is important to gain functional relevance to natural contexts for the individual being treated.

Motor aspects of manual cueing. Manual cueing evidenced in the literature, as in Jordan's gestures (Square, 1999), has focused primarily on visual cueing, which is a definite aspect of the manual mimicry cue. The evidence for visual cues in the literature is strong (i.e., EPG, ultrasound, spectrograms, EMA, etc.); however, another potential contribution to the success of manual mimicry cueing lies in the motoric realm. The use of the manual mimicry cue is supported by evidence from traditional articulation treatment findings (i.e., kinesthetic-tactile cueing, etc.); however, manual mimicry is also supported by other theoretical perspectives. For example, Smith, McFarland, and Weber's (1986) finger tapping study concluded that the speech system is affected by the hand system. This current study is consistent with Smith and colleagues' finding of the coordination, or entrainment, across those two systems. The findings of this current study are also consistent with the theoretical model of dynamic systems. According to this view, the manual movement increased the accuracy of speech production by capitalizing on the interconnection between the two systems.

The use of the manual mimicry cue has theoretical implications consistent with findings of the link between the two dynamic human systems of speech and manual movement. The literature that examines the link between speech and manual movement tends to focus on linguistically-based aspects (i.e., how gesture is used to communicate); However, in this study the close tie between these two systems appeared to be intrinsically motor-based (i.e., how the hand movement affects speech production).

Another theoretical link from the results of this study to dynamic systems theory occurred because the participant began to spontaneously self-cue to improve her /r/ productions. It appears that this manual movement may be capable of coordinating and integrating information across the manual and speech systems. Iverson and Thelen (1999) discussed the tight synchrony of these two systems [speech system and hand movement] in terms of communicative gestures; however it may be of value to examine this synchrony of speech and manual movements solely within the motoric realm.

This connection between the manual and speech systems deserves further investigation based on the preliminary results that the link may facilitate learning of new speech motor movement patterns. The current study does not propose to determine the underlying mechanism for the connection between these two systems or across the function of these two systems, leaving the determination of these connections for future research.

The efficacy of manual mimicry could also be explained with respect to the neurosensory motor system, as discussed by van der Merwe (1997). She explains that when the core motor plan execution results in a deviation from accuracy (i.e., articulation disorders) there is a reliance on sensory feedback loops to attempt to correct errors in production. These sensory feedback loops were utilized in the current study to supplement the participant's external feedback (i.e., visual and kinesthetic) to increase correct productions. Manual mimicry was used to depict the force and spatio-temporal relationships of oral articulators.

Specific to the principles of motor learning, this study assisted the participant with developing internal understanding and representation of correct versus incorrect productions, given knowledge of results on a randomized schedule. This internalization was evidenced by the participant's spontaneous use of the manual mimicry cue to self-correct erred productions.

Regarding resource allocation theory, this particular treatment appears to have been effective in using the concept of hands to make memories (Cook, Yi Yip, & Goldin-Meadow, 2010). The use of the hands possibly eased the burden on the auditory and sensory feedback loops that were required for subtle adjustments by the participant in her internal development of correct versus incorrect productions, as well as decreasing the overall cognitive load during new learning.

The results of manual mimicry treatment, as detailed in this study, can be explained by several links to the theoretical perspectives: entrainment of motor systems, dynamic systems theory, neuromotor sensory framework, the principles of motor learning, and resource allocation theory. Manual mimicry cueing shows promising results for clinical treatment of intractable /r/.

As addressed throughout this document, the importance of finding evidence for an effective /r/ treatment cannot be understated. This particular study demonstrated exploratory single-subject level evidence to support the use of manual movement cueing in the treatment of /r/ distortions. The implication for clinicians treating /r/ (especially in school settings) is that they have access to a novel treatment tool with single-subject level evidence without excessive cost, time, or need for instrumentation. If implemented, it is suggested that clinicians take rigorous outcome data when utilizing manual mimicry to both evaluate its effectiveness with a particular individual and also as continued documentation of the overall efficacy of this treatment.

Limitations

Two major limitations of the current study were the need for a greater sample size and a potential clinician bias toward use of the manual mimicry cue based on prior clinical knowledge and experience. Secondary limitations were limited participant training at a variety of levels of speech sound production (deeper analysis is needed at word, sentence, and conversational level), examining generalization of treatment, and assessing maintenance.

Potential threats to internal validity in this study included the following. Participant maturation and effects of any unaccounted for external events that the participant may have experienced because the research and data collection lasted approximately one month. The subject selection may be a threat to validity due to convenience sampling. Most research has potential to demonstrate the Hawthorne effect, in which participants act differently than they would in natural contexts due to the controlled research environment. Although attempts were made to control for potential confounds during therapy by counterbalancing and randomizing stimuli, potential learning or test practice sequencing effects may have been present due to the use of controlled stimuli. Finally, another potential threat to validity was instrumentation or equipment measurement errors using the EPG. Although the instrumentation was tested and appeared functional, the results indicate a consistency despite perceptual change that suggests the possibility of equipment error.

Threats to external validity include: the sample size and subject selection, the findings may not be generalizable to the population of all people with /r/ distortions, especially due to an age restriction in the inclusion criteria in this study. Geographic sampling was limited with one participant. The participant's specific historical and demographic background may be important factors, limiting the generalizability of the results.

Future research directions. The first area for future research would be replication studies for manual mimicry training at the word and conversational levels to establish efficacy at the single-subject level by different investigators and for different levels of production. Replication studies are needed to verify reliability and treatment effect. It is suggested that future studies employ a much larger sample size to increase the statistical power of the data and that a more varied age range be included. The acquisition of data on typical /r/ productions is an area of potential research, considering the high variability found within and between speakers. Future research could also focus on alternate dependent variables including acoustic analysis of formant frequencies F1-F5. Other non-instrumental forms of cueing paired with manual movement cues should be explored, namely tactile-kinesthetic cueing techniques (i.e., Lynn Carahaly's Speech EZ Apraxia Program and the Speech Buddy for /r/ placement direction and kinesthetic feedback). Further exploratory studies of manual mimicry using the vowel quadrilateral for clinicians would be a potential avenue for research. Measurements of timing and synchronicity of manual movements during accurate /r/ productions in comparison to inaccurate productions could be completed. Although, not proposed in the current study, the details gleaned from the qualitative portion suggest that a more thorough qualitative analysis in this area would be beneficial. The data set from the current study could also be examined using descriptive statistics. A measure of tongue tension as opposed to placement would be an interesting avenue for future research to further explore that hypothesis.

Another avenue for future research is in the related area of motoric coordination of the speech production system and manual motoric system in the absence of linguistic intention. Basic science research regarding the connection or link across these two systems needs to be done to substantiate speech system entrainment with the hands as suggested by the findings from this study. A basic science understanding of the underlying mechanism of the connection between these two systems and across the function of these two systems is necessary. Lastly, future research could examine if self-cueing (internal feedback), as opposed to clinician driven (external feedback), increases generalization and maintenance of treatment effects.

Summary

In summary, this exploratory single-subject study shows promising results for remediating /r/ errors at the vocalic level by offering a tool, manual mimicry, with potential for increased efficacy (given future research), for clinicians without access to budgets, resources, and equipment. This study also has ramifications for the direction of motor entrainment research in the area of motor speech execution and its coordination with the manual system. Further validation of manual mimicry cueing is necessary at higher levels of evidence with larger sample sizes; however, this treatment had a functionally relevant effect on this young adult with an intractable /r/ error that had been previously not responded to treatment.

APPENDIX

Appendix A Stimuli of vocalic /r/ probes

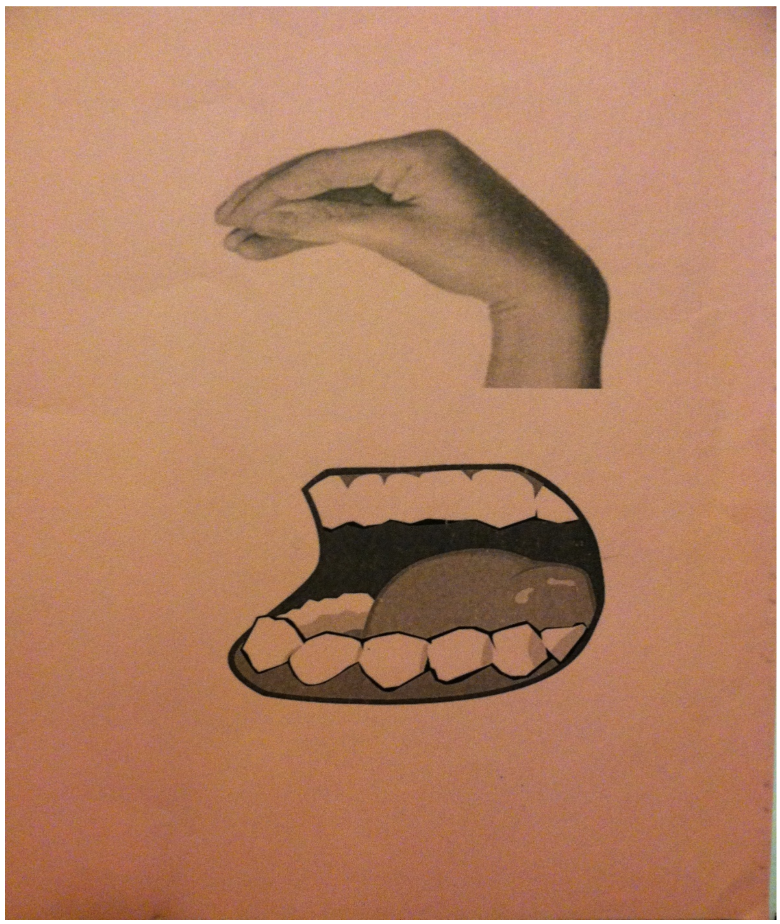
Presented on 2 x 4” card stock.

Stimuli lists of word level probes

Baseline1 presentations		Baseline2 presentations		CVC probes	Generalization probes
Internet	Wear	Share	Rope	Par	Rose
Organ	Scar	Drink	Racetrack	Mar	Sister
Wire	Year	Rugby	Entrance	Har	Pear
Wrap	Berry	Iron	First	Hire	Story
Raw	Artery	Burger	Crust	Hair	First
Fort		Siren	Front	Bare	Rope
Sailor		Cereal	Story	Bore	Hungry
Prize		Farm	Fries	Core	Red
Grape		Steer	Pretzel	Gore	Relaxed
Sore		Army	Recline		Stare
Where		Earn			Portrait
Finger		Area			Turn
Roast		Refrigerator			Pretzel
Very		Author			Fries
Bar		Beret			Perfect
Tree		Poor			Raindrop
Bread		Trap			
Mark		Rye			
Hairy		Hungry			
Perch		Swear			

Presented on 2 x 4” card stock.

Appendix B Graphic illustration of hand cue representing oral articulators



Presented on 8 ½ x 11" letter size card stock.

Appendix C Session Structure

Session	Length *(in mins.)*	Assessment/Tx	Operationalized details	Length of time/ task *(mins.)*	Rationale	Probe/ DV	Total probes	Probe type
pre		informed consent, assent, permission documents and case history form			informed consent			
		Palate mold created	orthodontist appointment		required for EPG use			
1	60	warm-up		5		conversation al speech sample		
		/r/ baseline #1	contextual baseline /r/ probe- 25 words	15	/r/ all contexts word level	25 randomized /r/ words	25	Bcx_1
		OME	Oral motor mechanism exam (adapted from Robbins & Klee)	5	assessment of oral mechanism			
		Peabody Picture Vocabulary Test	PPVT-4	10	assessment of language			
		break		10				
		Hearing/Vision Screening	Audiometer, tympanometry/ visual probe list	5	assessment of hearing and vision			
		/r/ baseline #2	baseline /r/ probe- 25 vocalic /r/	10	vocalic /r/	25 randomized vocalic /r/	25	B_2
2	60	Palate fit	warm-up	3	equipment test/ conversational sample			
		/r/ baseline #3	baseline /r/ probe- 25 vocalic /r/	10	3rd baseline w/ EPG	25 randomized vocalic /r/	25	B_3 EPG_A
		Goldman-Fristoe Test of Articulation	GFTA-2	10	assessment of articulation			
			break	10				
		TONI	Test of Non Verbal Intelligence	10	assessment of cognitive function			
		manual mimicry (MM) therapy (tx)	see MM tx details	9	manual gestures			
		MM probe	/r/ probe- 25 vocalic /r/	8		25 randomized vocalic /r/	25	MM_1
3	60	Palate fit	warm-up	3	equip test/ conver. sample/ continual contextual baseline	3 continous contextual baseline /r/ word probes	3	B_{c1}
		MM tx		15				
		MM probe		8		25 randomized vocalic /r/	25	MM_2
		break		10				
		MM tx		15				
		MM probe	/r/ probe- 25 vocalic /r/	9		25 randomized vocalic /r/	25	MM_3
4	60	warm-up	warm-up/conversation/contextual baseline	3	equip test/ conver. sample/ continual contextual baseline	3 continuous contextual baseline /r/ word probes	3	B_{c2}
		MM tx		15				

Appendix D Recruitment Flyer

PARTICIPANTS NEEDED

Individuals between the ages of 12 and 30 are needed to participate in a research study looking at treatment for /r/ sound errors.

Requirements:

/r/ sound error as primary speech sound disorder
English spoken as primary language
No history of language or hearing problems
No orthodontia currently
Participation will require nine 60-minute visits to Fisher Hall at the Duquesne University and one visit to a local orthodontist to have a mold of your mouth created

For more information, please call or Email

Jessica Lynch

412-606-1477

lynchj4@duq.edu

Speech Study 412-606-1477 lynchjessica1@gmail.com	*Speech Study 412-606-1477 lynchjessica1@gmail.com*	*Speech Study 412-606-1477 lynchjessica1@gmail.com*	*Speech Study 412-606-1477 lynchjessica1@gmail.com*	*Speech Study 412-606-1477 lynchjessica1@gmail.com*	*Speech Study 412-606-1477 lynchjessica1@gmail.com*	*Speech Study 412-606-1477 lynchjessica1@gmail.com*

Appendix E Phone Screening Script

Thank you for your interest in participating in our study at the Duquesne University Speech-Language-Hearing Clinic for individuals ages 12-30 to examine the effects of manual gesture treatment on /r/ speech sound errors. The overall objective of this study is to observe the effect of gesture use for treating /r/ errors. Individuals eligible for the study include:
Individuals between the ages of 12 and 30 years old, who have been diagnosed with a speech sound disorder including incorrect production of the /r/ sound.
Must be free of orthodontia and oral prosthetics
No additional diagnosis of a language or hearing disorder

If you are eligible and agree to participate in this study, your child will receive speech, language, and hearing evaluations which will require nine 1hour sessions at the Duquesne University Speech, Language, and Hearing Clinic. During the sessions your child will interact with a speech-language pathologist who will administer speech, language, and hearing tests, an assessment of your oral motor abilities, and a test of non-verbal intelligence.
1) How old are you?
2) Have you had a normal development thus far? Significant birth history? History of ear infections?
3) How would you describe your speech?
4) Do you experience errors with the /r/ sound in their speech? If so, about how long has this been occurring?
If so, have you received treatment to address a speech sound disorder?
Describe the frequency, duration, location, and other relevant information about their treatment.
5) Have you ever been treated for other speech sound errors or language issues? If so, what were you treated for? How long were you treated? Are you still being treated?
6) Do you have any motor, neuromotor, or behavioral problems that you are concerned about? Have you ever been seen by an OT or PT?
8) Are you a native English speaker? Are there any other languages spoken in the home?
9.) Do you have any orthodontia (i.e., braces, head gear…) or oral prosthesis (i.e., a palatal lift…)?
10.) Do you have functional vision and reading skills?
Are you able to read single words in type font size of 12 on a page held within 2 feet from you?

Thank you so much again for your interest in our study. Based on the answers you provided it appears at this time you do/do not qualify for our study.

Appendix F All context /r/ word examples

VOCALIC /r/		CONSONANTAL /r/			CONSONANT + /r/ CLUSTERS									
OPEN	**CLOSED**	**INITIAL**	**MEDIAL**	**FINAL**	**["th"r]**	**["sh"r]**	**[spr]**	**[rm]**	**[rt"sh"]**	**[kr]**	**[str]**	**[pr]**	**[rt]**	**[rd]**
fur	arm	read	story	stair	three	shrew	spring	arm	arch	crab	struck	prize	art	beard
fear	wrestling	creek	gary	spear	throw	shredded	spray	farm	march	crib	strum	price	court	blackboard
fare	pirate	rain	earring	airport	thrill	shrug	sprite	storm	porch	crack	strike out	prince	port	lord
tiger	girl	root	eery	starve	throttle	shrimp	sprinkles	worm	starch	cradle	strong	preschool	head start	flashcard
motor	growling	rat	teary	mark	thrift shop	shrub	spray gun	snowstorm	torch	crown	stroller	preach	heart	hard
finger	bird	road	perry	bear	thread	shroud	spruce	**[rn]**	research	crawl	street	prairie	tart	bored
grocer	shirt	reak	marry	deer	throat	shrivel	spry	born	**[tr]**	crumbs	stranger	prowl	starve	hoard
racer	burn	relax	barrel	bare	threw	shriek	sprint	corn	trip	crow	straw	pray	**[rk]**	award
sister	hurl	relaxed	europe	dear	throw rug	shred	sprinkle	horn	track	cricket	stronghold	proof	arc	**[rg]**
monster	nurse	rattle	harry	for	thriller	shrill	springboard	thorn	train	crete	stripe	prison	mark	morgue
chester	curl	robot	arrest	car	throne	shrink	sprinkling	torn	trent	christmas	string	press	bark	borg
her	hero	rug	irish	guitar	thrifty	shrine	springtime	popcorn	tractor	chris	strike zone	princess	pitchfork	**[rs]**
butter	ridge	rainy	zorro	door	throwing	shrewd	sprain	acorn	tree	croak	streetcar	prune	denmark	horse
other	front	red sea	hero	poor	threshold	mushroom	sprinkler		trunk	creak	stream	prick	ark	hoarse
flavor	run	raven	garage	jar	throng		spread		tray	crossbow	strap	pretzel	dark	remorse

Example /r/ word list from: Bleile, K. (2006). The Late Eight. San Diego, CA Plural Publishing.

Vocalic /r/														
AIR			**AR**			**EAR**			**IRE**			**OR**		
INITIAL	MEDIAL	FINAL	INITIAL	MEDIAL	FINAL	INITIAL	MEDIAL	FINAL	INITIAL	MEDIAL	FINAL	INITIAL	MEDIAL	FINAL
area	berries	ware	army	farm	scar	ear	cereal	year	iron	siren	wire	organ	fort	sore
prevocalic /r/		**blends**		**recurrent /r/**		**ER**								
						INITIAL	MEDIAL STRESSED		MEDIAL UNSTRESSED		FINAL			
wrap		prize		artery		earn	perch		internet		sailor			

Example vocalic /r/ word list from 21 types of vocalic /r/ and 11 blends by: Ristuccia, C., & McGovern, S. (2009). The Entire World of R Curriculum Book. Tybee Island, GA: Say It Right.

Appendix G Vowel Quadrilateral

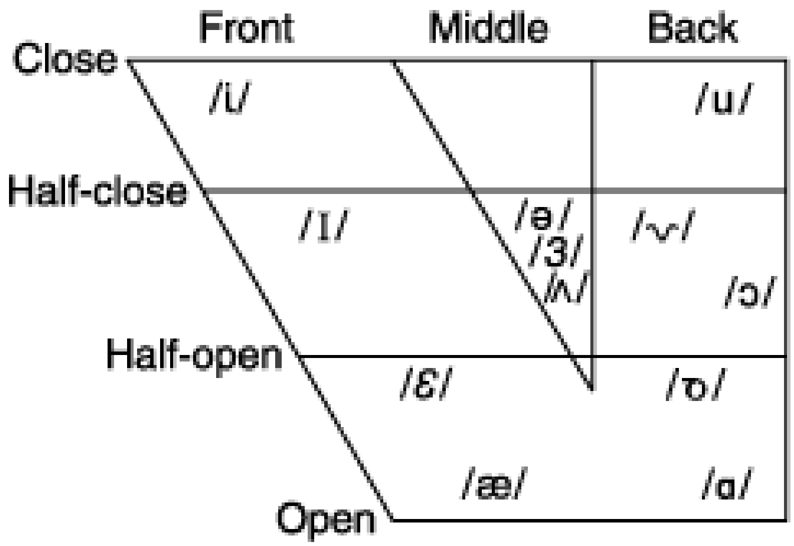


Image Retrieved from: https://engineering.purdue.edu/~ee649/notes/modeling.html

Purdue Electrical and Computer Engineering Course EE649: Speech Processing by Computer taught by Professor Leah H. Jamieson, School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN.

Appendix H Example data collection form for listeners

Naive listener form **Is the /r/ correct or incorrect?** Date: ___________

ex:	*Correct*	*Incorrect*	*target*	*difficulty*
a	*x*		*reed*	
b		*x*	*air*	*x*

Each target will be played twice. Each /r/ sound must be judged as correct or incorrect by putting an x in **either** the correct or incorrect box. You may indicate if you had difficulty judging correctness of a particular /r/ sound with an x in the difficulty box.

	Correct	**Incorrect**	target	*difficulty*		**Correct**	**Incorrect**	target	*difficulty*
1			oar		46			ire	
2			are		47			oar	
3			air		48			are	
4			army		49			ire	
5			are		50			air	
6			air		51			iron	
7			ear		52			air	
8			are		53			are	
9			air		54			oar	
10			ire		55			air	
11			ear		56			ear	
12			are		57			air	
13			are		58			ear	
14			ire		59			ire	
15			oar		60			are	
16			oar		61			oar	
17			oar		62			are	
18			ear		63			are	
19			air		64			are	
20			are		65			air	
21			ire		66			air	
22			ear		67			oar	
23			rare		68			air	
24			oar		69			ear	
25			are		70			air	
26			ire		71			ear	
27			air		72			oar	
28			trap		73			ear	
29			are		74			oar	
30			air		75			ire	
31			oar		76			air	
32			ear		77			oar	
33			ear		78			ire	
34			ire		79			oar	
35			ire		80			ear	
36			ear		81			are	
37			air		82			ear	
38			air		83			ire	
39			berry		84			ear	
40			are		85			ire	
41			ear		86			oar	
42			air		87			oar	
43			are		88			air	
44			ire		89			ear	
45			are		90			air	

Thank you very much for your participation!

Appendix I Transcriptions

	Clinician	Expert 1	Expert 2	Broad	Narrow	Notes
1	ɑɪɚ̮	ɑɪj ɚ̮	aɪ ɚ̮	1	0	addition medial j; dehroticized /r/ final position
2	i·ɚː	ɪ·r	ɪr	1	0	lengthening
3	r̮ʷeɪ r̮	wɛ r̮	r̮ɛ r̮	0	0	w/r initial substitution; dehroticized /r/ final position
4	ʊ r̮	ʊ r̮	ɔ r̮	1	1	dehroticized /r/ final position
5	ɑ r̮	wɛ r̮	ɑ r̮	0	0	addition initial w; dehroticized /r/ final position
6	ɑɪɚ	ɑɪɚ	aɪɚ	1	1	
7	ɛr	ɛr	ɛɚ	0	0	add schwa
8	træp	træp	træp	1	1	
9	ɑ r̮	ɑ r̮	a r̮	1	1	dehroticized /r/ final position
10	ɔrː	ɔrː	ɔr	1	0	lengthening
11	ɪ r̮	ij ɚ̮	ɪ r̮	0	0	addition medial j; add schwa; dehroticized /r/ final position
12	ɪr	ɪr	ɪr	1	1	
13	ɑɪ ɚ̮	ɑɪɚ	aɪɚ	1	1	
14	ɪr	ɪr	ɪr	1	1	
15	ɛrː	ɛr	ɛɚ	0	0	
16	ɛrː	ɛr	ɛɚ	0	0	
17	ɑr	ɑ r̮	ar	1	0	dehroticized /r/ final position
18	ɪ r̮	ijɚ	ɪ r̮	0	0	addition medial j; add schwa; dehroticized /r/ final position
19	ɛrː	ɛrː	ɛɚ	0	0	add schwa; lengthening
20	ɑ r̮	ɑ r̮	ar	1	0	dehroticized /r/ final position
21	ɑr	ɑr	ar	1	1	
22	ɑɪ ɚ̮	ɑɪ ɚ̮	ɑɪ ɚ̮	1	1	dehroticized schwa /r/ final position
23	ɔrː	ɔ r̮	ɔr̮	1	1	dehroticized /r/ final position
24	ɑr	ɑr	ar	1	1	
25	ɛrː	ɛr	ɛɚ	0	0	add schwa
26	ɑɪj ɚ̮n	ɑɪjɑn	aɪ ɚ̮n	0	0	addition medial j; add schwa; dehroticized schwa /r/ final position
27	ɛr̮	ɛr̮	ɛr	1	0	dehroticized /r/ final position
28	ɔr	ɔ r̮	ɔr	1	0	dehroticized /r/ final position
29	ɛr	ɛr̮	ɛr	1	0	dehroticized /r/ final position
30	ɪrː	ɪ r̮	ɪr	1	0	dehroticized /r/ final position
31	ɪ r̮	ɪr	ɪr	1	1	
32	ɑɪ ɚ̮	ɑɪr	aɪɚ	0	0	add schwa
33	ɑrː	ɑrː	ar	1	0	lengthening
34	ɑr̮	ɑr	ar	1	1	
35	ɑr̮	ɑr̮	ar̮	1	1	dehroticized /r/ final position
36	ɛr̮	ɛr̮	ɛr	1	0	dehroticized /r/ final position
37	ɛrː	ɛr̮ː	ɛr	1	0	dehroticized /r/ final position; lengthening
38	ɪɚ	ɪr	ɪr	1	1	
39	ɔr	ɔrː	ɔr	1	0	lengthening
40	ɪ r̮	ɪ r̮	ɪr	1	0	dehroticized /r/ final position
41	ɔ·rː	ɔrː	ɔr	1	0	lengthening
42	ɛrː	ɛr	ɛr	1	1	
43	ʌr̮	ɔ̠r̮	ɔr̮	1	0	retracted /ɔ/; dehroticized /r/ final position
44	h ɑ ɪ ɚ̮	h ɑ ɪ ɚ̮	aɪɚ	0	0	addition intial /h/; dehroticized schwa /r/ final position
45	ɔr̮	ɔr̮	ɔr	1	0	dehroticized /r/ final position
46	ɑr	ɑr	ar	1	1	
47	ɪr̮	ɪr̮	ɪr	1	0	dehroticized /r/ final position
48	ɑ ɪɚ	ɑ ɪɚ	aɪɚ	1	1	
49	ɪr	ɪr	ir	1	1	
50	ɑ ɪɚː	ɑ ɪɚ	aɪɚ	1	1	
51	ɔ·r	ɔ·r̮ː	ɔr	1	0	dehroticized /r/ final position; lengthening vowel + /r/
52	ɔrː	ɔr	ɔr	1	1	
53	ɑ ɪ ɚ̮	ɛr̮	ɛr	1	0	dehroticized /r/ final position

Appendix J Naïve Listener Observations

Observational themes of naïve listeners	
Difficulty	**Characteristics of the /r/**
This was a little difficult.	I couldn't tell if it was a /w/ sound or if there was a glide.
It was harder than I thought to tell correct/incorrect.	It was hard to judge "IRE" since it's not a real word that made it harder.
Listening to repetitous vocalic sounds makes you question your self "I don't know anymore".	"IRE" was the hardest for me to judge. I tried to think of how I would say it. I think she was adding the /j/ sound in there. I think it would be easier in a word.
/r/ is a weird sound.	"ORE" was the hardest for her to say- it was the most incorrect.
This made my head spin. I got lost.	I noticed "ORE" was the hardest sound for her to make.
This required a lot of listening and attention.	The production of "EAR" was almost always/consistently correct.
It was good to have the target sounds listed to follow along and know what they should be.	I noticed one production where she didn't glide the sound or make it liquid- she almost cut it off like a stop sound. That was a notable difference.
I treated /r/ before, so it wasn't challenging to judge.	**Elongation of productions**
Naïve listener judgment techniques	Was it ok for her to lengthen productions? Was she trained to do that?
I needed to develop an internal model of a correct /r/ because I was losing set and had to say /r/ in my head to compare to the productions.	It was harder with increased length to decide if the sound was right because she would hold it- hit it- distort it- then hit it again.
I was using self-practice, where I'd say the target silently then listen to the production again to judge against.	I had difficulty when sound was elongated. It was more questionable. It was easier to judge when the sound was faster.
Judgment was difficult when the same sound was back to back- I began to compare between productions (i.e., "ORE, ORE, ORE").	I could kind of tell when she started improving - that was when she started prolonging the productions.
I tried to use the "if I heard this on the street would it sound right to me" benchmark to help me judge.	With lengthier productions I would hear the sound derhoticized, but then she would fix it within that same production (difficult).
I caught myself saying it to figure out what she was doing.	**Emphasis, exagerration, and volume of productions**
I would compare them to each other if the same sound was put together a few in a row.	Some of the productions were louder than others and it seemed that the louder ones were more accurate.
I noticed improvement between productions because some were really bad and some were really better.	I noticed increased volume at times.
As a listener, your standards change with knowledge of this individual's /r/ issues, so you give credit for approximations.	When the sounds were louder there didn't seem to be as much effort and when there was emphasis on the sound those ones were easier to differentiate.
Having a target sound created a bias for me- I started looking at the target instead of just listening.	It was hard when sound was exagerrated.
I tried to score just for the /r/ and ignore the vowel.	Some /r/s were more exagerrated than others.
I was her previous clinician and I recognized improvement.	Some she had no emphasis and other she hit really hard. It seemed like the ones with more effort were better.
Seeing the person may have helped judge sound better.	**Vowel distortions in vocalic productions**
Vocalic vs. word level productions	It was harder and more distracting when she had difficulty with the vowel sounds
The initial position in the word level seemed more difficult, and medial and final seemed easier for her.	Her vowels were off and that made it hard- some were close and some were distorted.
/r/ initial words seemed hard for her.	**Tremor (vocal)**
The word level was easier to distinguish than the syllables.	It sounded laryngeal, in her throat, was that why she was so shaky?
At the word level, her performance decreased.	Her voice was so shaky it made it hard.
At the word level, she seemed to produce them faster and louder.	I noticed a tremor. Could it be that she was working her oral muscles so hard that they started to shake?
	Naturalness
	When she held the sound it was harder to decide if the sound was right. She was making the sound correctly, but it was not natural or normal sounding.
	She had these hard onsets that dropped off at the end and some had severe derhotization, almost like deaf speech.
	The productions were unnatural sounding.

References

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

Alwan, A., Narayanan, S., Haker, K. (1996). Toward articulatory-acoustic models for liquid approximants based on MRI and EPG data. Part II. The rhotics. *Journal of Acoustical Society of America, 101 (2),* 1078-1089.

American Speech-Language-Hearing Association. (2007). *Childhood Apraxia of Speech* [Technical Report]. Available from www.asha.org/policy.

Brown, L., Sherbenou, R.J., & Johnsen, S.K. (2010). *Test of Non-Verbal Intelligence, Fourth Edition.* San Antonio, TX: Pearson.

Byun, T.M. & Hitchcock, E.R. (2011). Treating /r/ Misarticulation. Session presented at the American Speech-Language-Hearing Association Annual Convention, Philadelphia, PA, November.

Carahaly, L. (2012). Overview of the Speech EZ Apraxia Program. Session presented at the National Conference on Childhood Apraxia of Speech Sponsored by CASANA. Boston, MA. Retrieved from: http://www.speech-ez.com.html.

Clark, C.E., Schwarz, I.E., & Blakeley, R.W. (1993). The removable R-appliance as a practice device to facilitate correct production of /r/. *American Journal of Speech-Language Pathology, 2,* 84-92.

Clayton, M., Sager, R., & Will, U. (2004). In time with the music: The concept of entrainment and its significance for ethnomusicology. *ESEM Counterpoint 1,* 1–82.

Complete Speech (2011). Available from http://completespeech.com

Cook, S.W., & Goldin-Meadow, S. (2006). The role of gesture in learning: Do children use their hands to change their minds? *Journal of Cognition and Development, 7(2),* 211-232.

Cook, S.W., Yi Yip, T.K., & Goldin-Meadow, S. (2010). Gesturing makes memories that last. *Journal of Memory and Language, 63,* 465-475.

Crow Hall, B.J. (1991). Attitudes of fourth and sixth graders toward peers with mild articulation disorders. *Language, Speech, and Hearing Services in Schools, 22,* 334-340.

Dagenais, P. (1995). Electropalatography in the treatment of articulation/phonological disorders. *Journal of Communication Disorders, 28,* 303-329.

DeThorne, L.S., Johnson, C.J., Walder, L., & Mahurin-Smith, J. (2009). When "Simon Says" doesn't work: Alternatives to imitation for facilitating early speech development. *American Journal of Speech-Language Pathology, 18,* 133-145.

Dunn, L. M., & Dunn, L. M. (1997). *Peabody Picture Vocabulary Test, Third Edition.* Circle Pines, MN: American Guidance Service.

Edeal, D.M., & Gildersleeve-Neumann, C.E. (2011). The importance of production frequency in therapy for childhood apraxia of speech. *American Journal of Speech-Language Pathology, 20,* 95-110.

Felesenfeld, S., Broen, P., & McGue, M. (1994). A 28-year follow-up of adults with a history of moderate phonological disorder: Educational and occupational results. *Journal of Speech and Hearing Research, 37,* 1341-1353.

Freed, D.B., Marshall, R.C., & Frazier, K.E. (1997). Long-term effectiveness of PROMPT treatment in severely apractic-aphasic speaker. *Aphasiology, 11,* 365-372.

Gast, D.L. (2010). *Single subject research methodology in behavioral sciences*. New York: Routlegde.

Garcia, J.M., Cannito, M.P., & Dagenais, P.A. (2000). Hand gestures: Perspectives and preliminary implications for adults with acquired dysarthria. *American Journal of Speech-Language Pathology, 9,* 107-115.

Gibbon, F. (2008). *The Handbook of Clinical Linguistics: Instrumental analysis of articulation in speech impairment.* Malden, MA: Blackwell Publishing.

Gibbon, F. (1999). Widening access to electropalatography for children with persistent sound system disorders. *American Journal of Speech-Language Pathology, 8,* 319-334.

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

Gibbon, F., Stewart, F., Hardcastle, W.J., & Crampin, L. (1999). Widening access to electropalatography for children with persistent sound system disorders. *American Journal of Speech-Language Pathology, 8,* 319-335.

Goldman, R. & Fristoe, M. (2000). *Goldman-Fristoe Test of Articulation, Second Edition.* Circle Pines, MN: AGS.

Goozee, J., Murdoch, B., Ozanne, A., Cheng, Y., Hill, A., & Gibbon, F. (2007). Lingual kinematics and coordination in speech-disordered children exhibiting differentiated versus undifferentiated lingual gestures. *International Journal of Language Communication Disorders, 42,* 703-724.

Hardcastle, W.J., Gibbon, F.E., & Jones, W. (1991). Visual display of tongue-palate contact; Electropalatography in the assessment and remediation of speech disorders. *British Journal of Disorders of Communication, 26,* 41-74.

Haynes, W.O., & Pindzola, R.H. (2008). *Diagnosis and Evaluation in Speech Pathology (7th ed.).* Boston, MA: Allyn and Bacon.

Howard, S. (2007). The interplay between articulation and prosody in children with impaired speech; Observations from electropalatographic and perceptual analysis. *Advances in Speech-Language Pathology, 9 (1),* 20-35.

Hustad, K.C., Garcia, J.M. (2005). Aided and unaided speech supplemental strategies: Effect of alphabetic cues and iconic hand gestures on dysarthric speech. *Journal of Speech-Language Pathology, 48,* 996-1012.

Iverson, J.M. & Thelen, E. (1999). Hand, mouth and brain: The dynamic emergence of speech and gesture. *Journal of Consciousness Studies, 6,* 19-40.

Kelly, S.D., Manning, S.M., & Rodak, S. (2008). Gesture gives a hand to language and learning: Perspectives from cognitive neuroscience, developmental psychology, and education. *Language and Linguistics Compass, 2,* 1-20.

Kratochwill, T.R., Hitchcock, J., Horner, R.H., Levin, J. R., Odom, S.L., Rindskopf, D.M., & Shadish, W.R. (2010). Single-case designs technical documentation. Retrieved from What Works Clearinghouse website: http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf.

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

McDonald, E. T., (1964a). *A Deep Test of Articulation. Pittsburgh: Stanwix House.*

McNeil, M.R. (1997). *Clinical management of sensorimotor speech disorders.* New York, NY: Thieme.

McNeil, M.R., Odell, K., Tseng, C.H. (1991). Toward the integration of resource allocation into a general theory of aphasia. *Clinical Aphasiology, 20,* 21-39.

McSweeny, J.L., & Shriberg, L.D. (1995). Segmental and Suprasegmental Transcription reliability. Phonology Project Technical Report #2, Waisman Center on Mental Retardation and Human Development, University of Wisconsin-Madison, 1-65. Retrieved from http://www.waisman.wisc.edu/phonology/techreports/TREP02.PDF

Mowrer, D.E. (1971). Transfer of training in articulation therapy. *Journal of Speech and Hearing Disorders, 36,* 427-446.

Peña-Brooks, A. & Hedge, M.N. (2007). *Assessment and treatment of articulation and phonological disorders in children* (2nd ed.). Austin, TX: Pro-ed.

Phillips-Silver, J., Aktipis, C.A., & Bryant, G.A. (2010). The ecology of entrainment: Foundations of coordinated rhythmic movement. *Music perception, 28,* 3-14.

Ristuccia, C. (2002). *Phonologic Strategy for /r/ Remediation.* ADVANCE for Speech-Language Pathologists & Audiologists, *12 (39),* 21.

Ristuccia, C. & McGovern, S. (2009). *The Entire World of R Curriculum Book.* Tybee Island, Georgia: Say It Right.

Rose, M. & Douglas, J. (2006). A comparison of verbal and gestural treatments for a word production deficit resulting from acquired apraxia of speech. *Aphasiology, 20,* 1186-1209.

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

Rusiewicz, H.L. (2010). The role of prosodic stress and speech perturbation on the temporal synchronization of speech and deictic gestures. Unpublished Ph.D. Dissertation, Pittsburgh University.

Rvachew, S. & Bernhardt, B.M. (2010). Clinical implications for dynamic systems theory for phonological development. *American Journal of Speech-Language Pathology, 19,* 34-50.

Sander, E. K. (1972). When are speech sounds learned? *Journal of Speech and Hearing Disorders, 37,* 55-63.

Shriberg, L.D. (1971). A response evocation program for /□/. *Journal of Speech and Hearing Disorders, 40,* 93-105.

Shriberg, L.D., & Kent, R.D. (2003). *Clinical Phonetics (3rd ed.).* Boston, MA: Allyn and Bacon.

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

Silverman, F.H. & Paulus, P.C. (1989). Peer reactions to teenagers who substitute /w/ for /r/. *Language, Speech, and Hearing Services in Schools, 20*, 219-221.

Smith, A., McFarland, D.H., & Weber, C.M. (1986). Interactions between speech and finger movements: An exploration of the dynamic pattern perspective. *Journal of Speech and Hearing Research, 29,* 471-480.

Spriestersbach, D.C., & Curtis, J.F. (2003). Misarticulation and discrimination of speech sounds. *The Quarterly Journal of Speech,* 484-491.

Square, P.A. (1999). Treatment of developmental apraxia of speech: Tactile-kinesthetic, rhythmic, and gestural approaches. In Caruso, A., & Strand, E. (Eds.) *Clinical management of motor speech disorders in children* (149-259). New York, NY: Thieme.

van der Merwe, A. (1997). A theoretical framework for the characterization of pathological speech sensorimotor control. In M. McNeil (Ed.), *Clinical Management of Sensorimotor Speech Disorders. New* York, NY: Thieme Medical Publishers, Inc.

Zhou, X., Epsy-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/. *Journal of Acoustical Society of America, 123*(6), 4466-4481.