USING ULTRASOUND IN SOUND PRODUCTION TREATMENT FOR ACQUIRED APRAXIA OF SPEECH: A CASE STUDY OF MULTIPLE SPEECH SOUND TARGETS

by

Winifred Murphey

B.A., The University of British Columbia, 2007

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Audiology and Speech Sciences)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

September 2016

© Winifred Murphey, 2016

Abstract

Acquired apraxia of speech (AOS) is a disorder of speech production. Treatment approaches and advancements are limited. Technological approaches or additions to current treatment approaches may provide greater benefit to individuals with AOS.

This case study examines the effects of treating AOS using an adapted Sound Production Treatment (SPT) hierarchy (Bailey, Eatchel & Wambaugh, 2015) that includes an ultrasound visual feedback component. The goal was to quantify the effects of two speech sound treatment blocks, one using SPT and one using SPT plus Ultrasound (U/S) for a single participant with AOS, and to compare the outcomes from each treatment block with one another.

The number of articulatory mismatches for treated and untreated speech sounds were analyzed based on their similarity to the treated speech sounds in terms of phonologic and articulatory features that can be seen on ultrasound.

Treated speech sounds showed increases in articulatory accuracy when using SPT with or without ultrasound. This was also true for untreated speech sounds that were maximally phonologically related to the treated speech sounds. Untreated speech sounds that shared articulatory/phonological features visible on ultrasound with the treated speech sounds showed increases in accuracy only in the SPT plus U/S condition. Untreated speech sounds that were minimally phonologically related to the treated speech sounds and were not visible on ultrasound showed limited improvement in either condition

These findings suggest that including a visual feedback route in speech treatment for AOS can be used to induce transfer of learning effects to speech sound targets with visibly similar articulatory gesture as those being treated. Clinically, using ultrasound as visual

ii

biofeedback in an SPT treatment may be effective in promoting transfer of positive treatment effects to speech sound targets when trained and untrained targets share characteristics observable on ultrasound.

Preface

My contribution to this thesis was in the formulation of the research question, design of the study, providing treatment sessions, data collection, data analyses, and interpretation of the results. Dr. B. May Bernhardt provided guidance in designing the study, interpretation of the results, and preparation of chapters. Dr. Penelope Bacsfalvi provided treatment sessions, offered guidance in treatment and assessment using ultrasound and assisted in designing the study. Dr. Tami Howe provided assistance in the preparation of all thesis chapters. This study was reviewed and approved by the Behavioural Research Ethics Board of the University of British Columbia under the name "Using ultrasound as visual feedback in speech therapy". The ethics certificate number is H04-80948.

Table of Contents

Abstract	ii
Preface	iv
Table of Contents	V
List of Tables	ix
List of Figures	Х
List of Abbreviations	xi
Acknowledgements	xii
Dedication	xiii
Chapter 1: Introduction	1
1.1 Overview of the Study	1
1.2 Acquired Apraxia of Speech (AOS)	1
1.2.1 Intervention Approaches for AOS	
1.3 Theories of Speech Production	
1.3.1 Models of Motor Learning and Rehabilitation	
1.3.1.1 Schema Theory	
1.3.1.2 Directions Into Velocities of Articulators (DIVA) Model	
1.3.2 Multi-Sensory Speech Processing	
1.4 Sound Production Treatment (SPT)	
1.5 Visual Biofeedback in AOS Treatment	
1.6 Ultrasound (U/S) in Speech Treatment	
1.6.1 Ultrasound in AOS Treatment	
	V

1.7 The Current Study and Research Questions	
Chapter 2: Methods	29
2.1 Participant	
2.2 Initial Assessment	
2.2.1 Elicitation Probe	
2.2.1.1 Probe Elicitation Procedure	
2.3 General Treatment Design	
2.3.1 Treatment Program	
2.3.2 Determining Treatment Targets	
2.3.3 Presentation of Stimuli in Treatment	39
2.3.4 SPT Treatment Block	40
2.3.4.1 SPT Hierarchy	
2.3.5 SPT Plus U/S Treatment Block	
2.3.5.1 SPT Plus U/S Hierarchy	
2.3.6 Treatment Fidelity	44
2.4 Outcome Measures	
2.4.1 Trained Listener Sound Target Accuracy Measures	
2.4.2 Untrained Listener Whole Word Accuracy Measures	
2.4.3 Articulatory Place Accuracy Measures Observed on Ultrasound	
2.4.4 Perceived Articulatory Place Accuracy Measures	
2.5 Analysis	
Chapter 3: Results	51
3.1 Outcomes as Rated by Trained Listeners on Speech Sound Accuracy	51
	vi

3.1.1 SPT and SPT Plus U/S: Treated Speech Sounds	51
3.1.2 SPT: Untreated Maximally Phonologically Related Speech Sounds	52
3.1.3 SPT Plus U/S: Untreated Maximally Phonologically Related Speech Sounds	52
3.1.4 SPT: Untreated Speech Sounds Visually Salient on Ultrasound	53
3.1.5 SPT Plus U/S: Untreated Speech Sounds Visually Salient on Ultrasound	55
3.1.6 SPT: Untreated Minimally Related Speech Sounds	55
3.1.7 SPT Plus U/S: Untreated Minimally Related Speech Sounds	55
3.1.6 Overall Effects of SPT and SPT Plus U/S on Speech Sound Accuracy	56
3.2 Outcomes as Rated by Untrained Listeners on Whole Word Accuracy	57
3.3 Evaluation of Ultrasound Videos	58
3.4 Effects of SPT and SPT Plus U/S on Accuracy of Articulatory Place	61
Chapter 4: Discussion	65
4.1 Outcomes of Treatment as Related to Research Questions	65
4.1.1 Treated Speech Sounds	66
4.1.2 Untreated Speech Sounds	67
4.1.3. Whole Word Accuracy	
4.2.4 Place of Articulation	72
4.2 Additional Outcomes of Treatment	73
4.2.2 Maintenance and Overgeneralization in SPT	74
4.3 Qualitative Impressions of Treatment	75
4.3.1 Sound Production Treatment for AOS	75
4.3.2 Ultrasound in Treatment for AOS	76
4.3.2.1 Ultrasound in Sound Production Treatment	
	vii

4.4 Limitations of the Study	
4.5 Future Directions	
4.6 Conclusions and Clinical Implications	
References	85
Appendices	92
Appendix A Mid-Sagittal Ultrasound Image of the Tongue	
Appendix B Probe Word List	
Appendix C Treatment Hierarchies	
C.1 SPT Block Treatment Hierarchy	
C.2 SPT Plus U/S Block Treatment Hierarchy	
Appendix D Treatment Fidelity Checklists	
D.1 SPT Block Treatment Fidelity Checklist	
D.2 SPT Plus U/S Block Treatment Fidelity Checklist	
Appendix E Probe Word Elicitation Transcriptions	
E.1 Baseline	
E.2 Pre-SPT	
E.3 Post-SPT	
E.4 Pre-U/S	
E.5 Post-U/S	

List of Tables

Table 2-1 Results of the Apraxia Battery for Adults- 2	31
Table 2-2 Measures of Inter-Rater Reliability for Trained Listeners	46
Table 3-1 Summary of Articulatory Mismatches	56
Table 3-2 Percentage of Correct Productions for Initial Elicitation Attempt	57

List of Figures

Figure 3-2 Mismatches for Treated Sounds	54
Figure 3-3 Mismatches for Sounds Visually Salient on U/S	54
Figure 3-4 Mismatches for Maximally Phonologically Related Sounds	54
Figure 3-5 Mismatches for Minimally Related Sounds	54
Figure 3-5 Total Mismatches for Treated and Untreated Sound Targets	57
Figure 3-6 Percentage of Correct Productions for Initial Elicitation Attempt	58
Figure 3-7 Pre-SPT Observed Accuracy of Place	60
Figure 3-8 Pre-U/S Observed Accuracy of Place	60
Figure 3-9 Post-SPT Observed Accuracy of Place	60
Figure 3-10 Post-U/S Observed Accuracy of Place	60
Figure 3-7 Place Mismatches for Treated Sounds	64
Figure 3-8 Place Mismatches for Sounds Visually Salient on Ultrasound	64
Figure 3-9 Place Mismatches for Maximally Phonologically Related Sounds	64
Figure 3-10 Place Mismatches for Minimally Related Sounds	64

List of Abbreviations

- AOS apraxia of speech
- SPT Sound Production Treatment
- CVA cerebrovascular accident
- GMP generalized motor program
- DIVA Directions Into Velocities of Articulators
- TMS transcranial magnetic stimulation
- KR knowledge of results
- KP knowledge of performance
- EPG electropatatography
- EMA electromagnetic articulography
- ABA-2 Apraxia Battery for Adults-2
- WAB-R Western Aphasia Battery- Revised
- D-COME Dworkin-Culatta Oral Mechanism Exam
- SLP Speech-Language Pathologist
- SLA- speech-language assistant
- IPA International Phonetic Alphabet

Acknowledgements

This thesis would not have been possible without a great deal of help and support. First, I would like to extend my deepest gratitude to Dr. B. May Bernhardt for her much needed guidance and indefatigable support throughout this process. I am also grateful to Dr. Penelope Bacsfalvi for lending her clinical expertise and advice to this research. I would like to thank Dr. Tami Howe for her valuable insights and willingness to take a chance on this project. I would also like to recognize Yasmine Bia for graciously offering her time to provide practice sessions during treatment and everyone at the North Shore Stroke Recovery Centre who gave their time, space, and resources to this project. I am eternally grateful to the participant in this study and her husband for their unparalleled commitment and enthusiasm.

I owe many heart-felt thanks to the School of Audiology and Speech Sciences class of 2016 who consistently inspire and astound me. I am especially grateful to Liv Meriano, Matthew Kowalyk, Shelby Siroski and Thyra Driver for assisting with the analyses in this paper.

I would also like to acknowledge Wendy Duke and Dr. Donald Derrick, for their prompt and practical answers to the vaguest of questions, Dr. Murray Schellenberg, for his encouragement during difficult moments, and Dr. Bryan Gick, who initially fostered my interest in ultrasound and has helped to provide me with some of my most memorable learning experiences.

Dedication

TO MY MOM FOR ALWAYS BELIEVING IN ME. I LOVE YOU TO PIECES.

AND

TO JERIK FOR HIS STRENGTH AND SUPPORT. YEP, I FINISHED MY THESIS.

AND

TO PROFESSOR MERLIN CLAW WHO ALWAYS KNOWS WHEN TO TAKE A BREAK.

xiii

Chapter 1: Introduction

1.1 Overview of the Study

Acquired apraxia of speech (AOS) is a disorder of speech production that can be difficult to treat. The research base for effective treatments is limited despite a clinical need for more evidence-based approaches. This thesis presents a case study aimed at expanding the evidence for a currently used AOS treatment and investigates a potential improvement to the approach.

Chapter 1 first introduces the definition of AOS that was used for this study and outlines treatment approaches for this disorder of speech production. Next, theories of speech production that served as a basis for the treatment study are discussed. A review of the literature on both Sound Production Treatment and ultrasound visual feedback in speech treatment is then presented. Finally, the hypothesis and three main research questions are outlined in this chapter. Chapter 2 provides background information on the participant of the study and presents the results of the initial assessment. It also describes the overall treatment design including outcome measures, and time points for data collection. Chapter 3 describes the results in terms of both auditory perceptual and visual coding measures. In Chapter 4, a discussion of the effects of the treatments are presented with respect to the hypothesis and research questions that were posed. Subjective observations from treatment are considered and the limitations of the study are explored. Finally, future directions for research and clinical implications are suggested.

1.2 Acquired Apraxia of Speech (AOS)

As noted, the speech production condition of focus in this study is acquired apraxia of speech (AOS). AOS is a disorder of speech production that is typically caused by a cerebrovascular accident (CVA). Clinically, diagnosis of AOS tends to be based on a pattern of

observable speech characteristics (McNeil, Robin, & Schmidt, 2009; Wambaugh & Shuster, 2008). Although disagreement remains regarding the precise diagnostic criteria that define the condition, researchers generally agree that the characteristics of AOS include perceived speech sound distortions, sound substitutions and prosodic disruptions (McNeil et al., 2009; Ogar, Slama, Dronkers, Amici, & Gorno-Tempini, 2005; Peach, 2004; Wambaugh & Shuster, 2008). Prosodic disruptions include slow speech rate and inaccurate stress patterns (McNeil et al., 2009). Speech sound distortions maintain some recognizable features of the intended speech target but sound imprecise. Variations on voicing could be examples of distortions, as in /s/ being produced inconsistently as [z]. An example of a complete speech sound substitution might be a /t/ being produced as [k] or [g]. Additional features of AOS are also mentioned in the research literature, e.g., effortful groping, repetitive attempts at productions, and difficulty with initiation, but these characteristics are not present in every case (McNeil et al., 2009; Wambaugh & Shuster, 2008).

The underlying cause of perceived speech sound difficulties in AOS is still widely contested. While theoretical frameworks of speech production will be presented later, here I present an introduction to the debate about the nature of AOS. The most common view in the research literature, which comes from McNeil and colleagues (2009), claims that AOS is a phonetic-motoric disorder of speech production caused by an inefficiency in translating filled phonological frames into the necessary and intended motor movements of speech. This means that impairments to the sequencing and timing of articulators, rather than difficulty with sound selection, leads to perceived problems in the execution of the speech string at both the segmental and intersegmental levels (McNeil et al., 2009; Ogar et al., 2005; Peach, 2004).

This view of AOS is by no means held by all researchers. Zeigler (2002) explores an opposing notion that the observable symptoms of AOS could be explained by an inability to access stored phonetic representations. The features of AOS have also been explained using a dual-route hypothesis of speech encoding (Varley & Whiteside, 2001). The dual routes of this model include an automatic processing mechanism that assembles high frequency phonetic patterns and stores them in a mental syllabary while a second route of production is reserved for more low frequency phonetic patterns which are assembled from sub-syllabic units (Levelt & Wheeldon, 1994). In a dual-route theory of AOS, neurological damage causes impairments in the speech system's automatic processing mechanism and without access to the mental syllabary the required phonetic representations must then be assembled phone by phone or feature by feature (Varley & Whiteside, 2001).

Even among those who subscribe to the view of AOS as a motor programming disorder (McNeil et al., 2009), discussions about the specifics of this definition are ongoing. Questions persist over what constitutes a "motor program" (Ballard, Granier, & Robin, 2000; Schmidt & Lee, 1999) and at what stage of speech processing these programs are impaired (Zeigler, 2002). The assumption that AOS is a speech-specific condition has also been challenged and a conceptualization of AOS as a disorder impacting attentional resources needed for speech and non-speech motor tasks has been presented (Ballard, et al., 2000). As a final challenge, it has even been suggested that there may be more than one subtype of AOS (Ogar et al., 2005).

It seems there is no single explanation of AOS that is currently available to clinicians and researchers. Despite the lack of specificity surrounding the condition, the resulting deficits in speech production can severely impact a speakers' ability to communicate, and as a result, their

social participation. For these reasons, it is important that clinicians have effective treatment options for this condition and seek evidence-based improvements to treatments that are currently being implemented.

1.2.1 Intervention Approaches for AOS

As one might suspect, based on the problems with definition, differentiation, and diagnosis outlined above, there are a limited number of evidence-based AOS treatment approaches. However, a systematic review by Wambaugh, Duffy, McNeil, Robin, and Rogers, (2006b) provides a sample of AOS treatments which are divided into four main categories. These categories are articulatory-kinematic, intersystemic facilitation, rate and rhythm control, and alternative or augmentative communication. It is important to recognize, however, that these categorical distinctions are primarily for ease of conceptualization. The different treatment approaches are not mutually exclusive and treatments that combine approaches do exist. The two categories relevant to this study are articulatory-kinematic and intersystemic facilitation.

Articulatory-kinematic treatments are those that attempt to improve the movement and positioning of articulators and tend to be the most well-researched of the AOS treatment approaches (Ballard, Wambaugh, Duffy, Layfield, Maas, Mauszycki, & McNeil, 2015; Peach, 2004; Wambaugh, Duffy, McNeil, Robin, and Rogers, 2006a). Intersystemic facilitation approaches intend to make use of relatively intact systems to facilitate improvements to speech production (Wambaugh et al., 2006a), for example pairing hand movements or limb gestures with speech sounds (Wambaugh & Shuster, 2008). These treatment approaches have received less rigorous study than articulatory-kinematic treatments. However, the assumption that the speech system can utilize input from other sources provides potential opportunities for improving

AOS treatments in clinical settings. Making use of intersystemic facilitation in combination with techniques from other broad treatment categories could be a way to improve upon the effectiveness of various AOS treatments. For example, intersystemic facilitation has been combined with rate and rhythm control approaches (i.e., finger tapping with a metronome) to target speech pacing (Wambaugh & Mauszycki, 2008).

Determining the most effective treatment for AOS is an ongoing challenge. Few experiments comparing the effectiveness of various treatment options have been undertaken (Peach, 2004). This may be in part because treatment goals for AOS are generally aimed at improving communicative effectiveness and therefore based upon individual need (Ogar et al., 2008); a comparison of altogether different treatments for AOS may not be appropriate. However, comparing a currently available treatment option applied with and without an adjunctive intersystemic facilitative element of treatment could be of value in the search for more evidence-based and effective treatments for AOS. This type of proposed comparative examination is the topic for this treatment study. Ultrasound feedback was used as the element of intersystemic facilitation applied with Sound Production Treatment (SPT), an articulatorykinematic treatment approach. The theories on which this experiment are founded are outlined next.

1.3 Theories of Speech Production

It is generally believed that speech is the result of numerous transformations in the brain from abstract to highly specific representations and eventually perceivable sounds. This means that motor execution of speech is the end result of processes based in the broader systems of cognition and language. Levelt, Roelefs, and Meyer's (1999) theory of lexical access explores

these broader systems of language, word selection, sound sequencing, and their place within a network of cognition, attention, and overall resource capacity. Connectionist models describe a distributed nature of language processing and speech production (Dell & O'Seaghdha, 1992). This suspected distribution of processing combined with an unclear understanding of the systems that govern intact speech (Bernhardt, Stemberger, & Charest, 2010) make it difficult to isolate and explain speech disruptions caused by neurological damage. Among researchers, drawing neuroanatomical correlates to the observable symptoms of AOS is contentious (Ogar et al., 2005). It is possible that damage to the brain structures that govern speech will necessarily impact language, a possibility which is not contradicted by the high rate of aphasia co-occurring with AOS.

However, the inquiries driving research on AOS and motor speech systems are based on the belief that motor execution, rather than higher level language processing, is where the specific deficits of AOS arise. Furthermore, empirical evidence on AOS treatments including the one used in this study, SPT, has mostly been explained using motor learning theories of speech production. Two of these theories, schema theory and Directions Into Velocity of Articulators (DIVA) model, are discussed below. This discussion is followed by a brief review of literature on multi-sensory speech processing in order to explore the theoretical basis for applying an intersystemic facilitation approach to treatment.

1.3.1 Models of Motor Learning and Rehabilitation

Motor learning is not defined by any single model although it is often associated with schema theory which will be discussed first. In a broad sense, motor learning refers to the process by which one acquires the capability for performing purposeful skilled actions (Schmidt

& Lee, 1999 p. 264). One of the key tenets of motor learning theories is that successful motor learning will lead not only to acquisition of a motor behavior, but also a transfer of learning to related motor behaviors, (e.g., the same behavior in different contexts or a different behavior that shares similar movements) and a retention of learning after practice has stopped (Bislick, Weir, Spencer, Kendall, & Yorkston, 2012; Maas, Robin, Austermann, Freedman, Wulf, Ballard, & Schmidt, 2008; Schmidt & Lee, 1999). Neural plasticity is suggested to be at the heart of motor learning, but these mechanisms of neural change are far from understood (Seitz, Matyas, & Carey, 2008). Some suspected principles that enhance motor learning have been identified from research on non-speech motor control. Motor learning principles are based on the observed relationships between acquisition, transfer and/or retention of skills in specific conditions of practice (Schmidt & Lee, 1999). These conditions include practice type, schedule, and variability, stimulus type, presentation, and complexity, as well as feedback type and frequency (Schmidt & Lee, 1999). At present, many of these principles have been confirmed in non-speech motor learning research while far fewer have been explored in the context of speech (Maas et al., 2008).

There is a generally held assumption though, that speech motor actions are built and refined on non-speech motor actions and a significant portion of our understanding of motor speech production is built on the principles of non-speech motor control (Maas et al., 2008). For this reason, speech clinicians and researchers tend to apply motor learning principles to speech treatments and use models of motor control as a way to understand speech production deficits. There is, admittedly, some risk to relying on non-speech motor learning research to predict speech outcomes. First, researchers have yet to define the crucial similarities and differences

between speech and non-speech motor control (Ballard et al.; 2002; Maas, 2006; Zeigler, 2002). Principles of non-speech motor control may not apply in exactly the same way as speech-motor control. In the section below on schema theory we further discuss some reasons for this proposition. Second, the majority of motor learning research has been performed on healthy subjects (Maas et al., 2008). When lesions to brain tissue occur, the representations or brain circuits previously established for motor learning may no longer be accessible (Seitz et al., 2008). Adults with brain damage may not be able to use the same neural mechanisms that control learning in typical adults without brain damage. It is possible that functional relearning of motor skills occurs through newly established neural pathways (Seitz et al., 2008) or previously unused collateral neural connections. For both of these reasons, it is difficult to be sure which principles of motor learning necessarily apply in a rehabilitative context for speech.

Despite these limitations, clinicians and researchers continue to employ principles of motor learning in attempts to treat motor speech disorders (Bislick et al., 2012; Maas et al., 2008). However, only very few of the various motor learning principles have actually been studied in motor speech treatment. Further, in the treatment of AOS, few studies have explored motor learning principles with much methodological rigor (Maas et al., 2008). A systematic review performed by Bislick and colleagues (2012) found that a single study out of five yielded results that were robust enough to suggest that one particular motor learning principle holds for treating AOS. The principle in question concerns stimulus presentation and is discussed further in the next section on schema theory. So far, the little evidence available suggests that some motor learning principles derived from non-speech models can be applied to speech motor learning, although which ones and to what extent is still unknown.

Proposing a theory that can model intact and disordered motor systems while accounting for speech as well as non-speech findings is beyond the scope of this thesis. Rather, motor learning based theoretical models of AOS, despite their incomplete picture of speech production, were used to establish a framework for treatment. The two constructs, schema theory and the DIVA computational model, have been suggested to share some key features (Maas et al., 2008) and elements of both models were used to make predictions about treatment outcomes. They have both been used as theoretical frameworks to explain the nature of AOS.

1.3.1.1 Schema Theory

Schmidt (1975) proposed a schema theory of motor control in 1975 to try to account for the outcomes he observed in motor learning research. A motor response schema in Schmidt's theory, is an abstract memory representation for a skilled action (Schmidt & Lee, 1999). The theory supposes that learning to perform new actions is the result of the forming new schemas based on experiences and that these schemas serve as the basic motor plans for skilled actions. Long-term retention and generalization of learned motor skills depends on the quick and efficient access to motor schemas and the execution of a selected motor plan under a variety of conditions. Although schema theory was developed based on non-speech motor learning literature, it has been adopted by some speech researchers as a model for treating motor speech disorders (Ballard et al., 2000; Knock, Ballard, Robin & Schmidt, 2000; Maas et al., 2008; Wambaugh & Nessler, 2004).

The basic component of a motor plan is a generalized motor program (GMP) which is a stored representation for a motor movement (Schmidt, 1975). From a speech perspective a GMP might represent relative timing of articulators or direction of movement relating to a specified

place or manner of articulation (Maas et al., 2008). Before a motor plan is executed a GMP is assigned a set of parameters that define the specifics of the movement in that particular time and context (Schmidt, 1975). In speech these parameters may provide the exact timing and direction of motion based on the current state of the articulators (Ballard et al., 2000; Maas et al., 2008).

To correctly articulate a unit of speech the GMP and its parameters must be appropriate to the situation. This means that the system must have access to information about the current state and speed of the articulators as well as the expected outcomes of their intended articulation. To that end, it is further proposed that two states of memory are involved in executing and updating plans for motor actions. Recall memory is responsible for executing learned movements and recognition memory is responsible for evaluating immediate outcomes of those performed actions (Schmidt, 1975). Recall memory may be required to access a motor plan while recognition memory may be required to use output feedback to update the schema of that motor plan for future execution.

A schema theoretical perspective might explain AOS as the result of any number of inefficiencies within the proposed system. Wambaugh West and Doyle (1998) proposed that possible causes for groping and articulatory errors in AOS could include problems controlling the rate, range, or timing of articulatory movements. These observed symptoms could be the result of an inability to select or activate a GMP and appropriate parameters for a particular speech sound or syllable (Knock at al., 2000; Maas et al., 2008). A theory posited in Maas et al. (2008) suggests that AOS may be the result of an inability to use sensory feedback to update the parameters of a GMP; support for this perspective might be the observation that speakers with AOS tend to make repeated incorrect attempts at the same word. Schema theory might explain

this as the result of an inefficiency in recognition memory. Effective treatments of AOS in this model are suspected to affect access to or implementation of a generalized motor program and its appropriate parameters, thus strengthening the specifications of a schema for a target sound (Bislick et al., 2012; Katz, McNeil, & Garst, 2010). Schema theory supposes that sufficiently strengthening the schema of a speech motor movement through treatment should result in speech motor learning (Knock et al., 2000).

Some predictions about speech motor learning can be made using this theoretical model. Two of these predictions include (1) learned speech movements will generalize to untrained movements that share the same or a very similar GMP; and (2) gains made as a result of training will be maintained after the cessation of treatment (Bislick et al., 2012; Knock et al., 2000; Maas et al., 2008).

Of the small amount of clinical AOS treatment research that has attempted to test schema based motor learning principles, outcomes have proven to be difficult to interpret. Related to stimulus presentation, Knock and colleagues (2000) studied this principle in a single subject design with two participants with AOS. Based on non-speech motor learning principles, they expected the blocked presentation of training stimuli to be associated with better acquisition of treated speech sounds and random presentation to be associated with better retention but slower acquisition. They found improved retention for the random presentation as predicted but a high degree of variability in acquisition rate (Knock et al., 2000). Distribution of practice is another principle that has yet to be confirmed across modalities. Wambaugh, Nessler, Cameron, & Mauszycki (2013) explored this principle in a single subject multiple baseline treatment study with four adult participants with AOS. They found that dose frequency and practice schedule did

not appear to differentially impact the effect of Sound Production Treatment (Wambaugh et al., 2013). This finding is in opposition to non-speech findings where distributed practice compared with massed practice is a positive predictor of learning (Schmidt & Lee, 1999). Even dosage of treatment has yet to be confirmed in speech treatment studies for motor speech disorders (Maas et al., 2008). Overall, the empirical research on AOS treatments that have been founded on schema theory shows many gaps and inconsistencies (Bislick et al., 2012; Knock et al, 2000; Maas et al., 2008).

The schema model suffers from additional theoretical limitations. A basic unanswered question is how initial movement can be made before a schema exists for it (Schmidt & Lee, 1999). Further, there is a lack of specificity as to what types of experiential input can lead the system to update a GMP with new parameters (Schmidt & Lee, 1999). Relative to speech production, schema theory may not sufficiently account for the conceptual differences between speech and non-speech motor movements. Zeigler (2002) specifically points out that speech production has different feedback mechanisms than non-speech motor movements. Although both movements provide tactile-proprioceptive feedback, speech provides audio-imitative feedback whereas limb movement additionally provides visuo-imitative feedback. Speech itself may be based in a frame of reference that is acoustic rather than articulatory in nature (Zeigler, 2002).

Schema theory's greatest limitation, however, i.e. its vagueness, is perhaps its greatest benefit. The question of what constitutes a "motor program" is one of the most noticeable aspects of this ambiguity (Ballard, et al., 2000; Guenther, 2006; Maas et al., 2008). Knock and colleagues (2000) posit that a GMP could be likened to an individual phoneme. Other

researchers have suggested that a GMP could be associated with an articulatory gesture, perhaps even an articulatory manner or place (Maas et al., 2008). There is, as of yet, no consensus as to what aspect of a speech gesture encompasses a singular GMP. This leads to difficulty in choosing therapeutic targets which will have the greatest effect on functional speech outcomes since a different definition of a GMP will lead to different suspected treatment outcomes, particularly in terms of transfer of motor learning to similar movements. If a GMP is a phoneme and motor learning occurs, improvements made for trained speech sound targets will transfer to those same treated targets in untreated contexts (e.g., training the speech sound "t" in the word "top" would improve articulation of the "t" in "hat"). If a GMP is an articulatory gesture and motor learning occurs, improvements made for trained speech sound targets will transfer to untreated speech sounds that share that articulatory gesture with the trained speech sound. For instance, if a GMP is an articulatory place gesture then training the speech sound "t" would improve articulation of the untreated speech sound "d" since they share articulatory place. From a theoretical point of view, without a firm assertion as to what components of speech make up a GMP, it is not clear what empirical findings would constitute transference of learning. Nevertheless, considering repeated execution of a parameterized GMP as the basis for motor learning may complement more explicit models of speech production such as the DIVA model.

1.3.1.2 Directions Into Velocities of Articulators (DIVA) Model

The DIVA Model is a neurobiologically grounded computational approach that aims to model the mechanisms of motor speech production and learning (Guenther, 2006; Tourville & Guenther, 2011). As a computational model, DIVA can theoretically be programmed to test predictions about speech motor learning in both healthy and disordered neurological systems.

However, in this thesis it is used only as a theoretical construct for hypothesizing about treatment outcomes.

The model assumes complex associations in the brain between stored representations of speech sounds and the motor and sensory information for planning a speech string (Guenther, 2006). In this model, speech sound production is controlled by a feedforward system, and a feedback system. The feedback system is further divided into the auditory and somatosensory subsystems. The model also proposes theoretical articulatory and somato-sensory error maps that are involved in learning to produce and refine the feedforward commands to articulators (Guenther, 2006; Tourville & Guenther, 2011). A learner depends on the feedback system initially to establish proficiency of speech motor movements and once sufficiently skilled, the brain relies on the feedforward system to operate more automatically (Guenther, 2006). The feedback mechanisms are assumed to have a major impact on the ability of a speaker to learn articulatory gestures and their auditory consequences. During speech learning, the feedforward commands are updated based on the sensory information being provided by the feedback systems. This sensory information is used to refine the articulatory error maps responsible for encoding the sensory expectations of articulations (Guenther, 2006).

The process of speech sound error mapping defined in the DIVA model appears to resolve some of the questions around how a speech motor plan arises prior to the development of strong feedforward representations, one of the limitations of schema theory. Another benefit of the DIVA model is that it offers two specific sources of sensory input that may influence motor speech learning. These sources of sensory information and (as it will be argued next) perhaps

others, may be part of the memory subsystems proposed by schema theory that help drive the selection and parameterization of GMPs.

Guenther (2006) also hypothesized that distortions in the system (e.g., warping of feedback) may limit the functionality of the automatic feedforward processes. Limited feedforward control might require that a learner rely more heavily on the feedback systems, to provide information about the sound and sensation resulting from each movement, in order to hone articulatory targets. Neurological damage to the feedback or feedforward systems of speech production could lead to an inability to develop, execute, or correct the necessary motor commands for speech (Tourville & Guenther, 2011). An explanation of the clinical characteristics of AOS based on DIVA has been proposed using the feedforward hypothesis (Maas, Mailend, & Guenther, 2015). The feedforward hypothesis maintains that AOS reflects a disruption of the feedforward control system of speech resulting in a stronger reliance on sensory feedback in order to produce selected speech sounds (Maas et al., 2015). Although the empirical evidence for this theory is still quite limited, it does suggest that a treatment approach that engaged spared feedback systems might improve treatment outcomes, for example treatments that were considered intersystemic in nature. Although the DIVA model feedback systems are composed of auditory and somato-sensory subsystems only, the introduction of visual feedback might also be able to influence motor speech learning.

1.3.2 Multi-Sensory Speech Processing

To justify using visual feedback as an intersystemic facilitation in AOS treatment it must be assumed that various types of multi-sensory input can affect speech processing. Hickok, Houde and Rong (2011) asserted the basic notion that auditory perception is involved in the

process of speech production and that the motor system is involved in speech perception. The computational model outlined by Hickok and colleagues (2011) will not be explored in detail here; however, this concept of a bilateral flow of sensory and motor information is basic to their notion of sensori-motor integration in speech. As noted above, the DIVA model considers that auditory and somato-sensory feedback are likely to play a role in speech production and speech motor learning (Guenther, 2006; Maas et al., 2015; Tourville & Guenther, 2011). Although the integration of visually presented feedback is not part of DIVA's feedback systems it is possible that other sensory inputs also influence speech production. Research on multi-sensory speech processing provides some support for this viewpoint.

Visual (McGurk & MacDonald, 1976) and even aerotactile information (Derrick & Gick, 2013) have been found to affect speech perception. More specifically, mismatches between visual or tactile input and acoustically presented speech can lead to misperceptions in speech sounds. McGurk and MacDonald (1976) found that subjects who heard the syllable "ba" while viewing the speech movements of the syllable "ga" tended to perceive the spoken syllable as "da". Derrick and Gick (2013) observed a similar phenomenon wherein subjects who heard unaspirated speech sounds (e.g., "ba" or "da") that were paired with air puffs presented to the skin were less accurate in identifying the acoustically presented speech sound than when no air puff was presented. The idea that cerebral motor systems are impacted by speech perception has also gained some support. For example, transcranial magnetic stimulation (TMS) studies (Fadiga, Craighero, Buccino, & Rizzolatti, 2002; Waktins, Strafella, & Paus, 2003) have found that hearing speech sounds can excite the cortical correlates assumed to be associated with tongue muscles required to produce the perceived speech sound (Fadiga et al., 2002). In addition,

visually presented speech movements excited the cortical correlates assumed to be associated with the observed lip muscles (i.e. orbicularis oris) (Watkins et al., 2003).

The idea of a bilateral flow between motor and auditory systems may not capture all possible sensory influences on the speech processing system. The perceptual system may attempt to make use of sensory stimuli provided despite the modality (Derrick & Gick, 2013). Sensory inputs that are not included in theoretical models of speech perception or production may exert their influence only when the input is made available.

Speech perceptual processing is strongly suspected to be a multimodal process and we are still learning the extent to which sensory and motor systems interact. No cohesive model of speech production that includes all sensory modalities is available. Further, how the speech production system might use visual feedback is speculative. Still, the evidence for multi-sensory processing in speech is sufficient to justify employing a speech treatment approach that attempts to exploit these multi-sensory systems.

The empirical evidence specific to the use of ultrasound as a visual feedback treatment tool is presented later in this introduction. First, we turn our attention to the AOS treatment that provided the foundations for this study that was adapted to include this ultrasound component.

1.4 Sound Production Treatment (SPT)

Sound Production Treatment (SPT) is the articulatory-kinematic treatment approach that was used as the basis for this study. SPT is a treatment approach that uses repeated practice within a response-contingent hierarchy to improve the production of targeted speech sounds consistently produced in error. SPT is assumed to be consistent with motor learning principles (Bislick et al., 2012; Wambaugh & Mauszycki, 2010; Wambaugh & Nessler, 2004). Specifically,

response feedback, repeated practice, and integral stimulation (i.e. "watch me, listen to me, say it with me") are assumed to promote speech motor skills in people with AOS (Wambaugh, West, & Doyle, 1998). The techniques used in all reported SPT hierarchies include modeling-repetition, orthographic cueing, integral stimulation, and articulatory placement cueing (Bailey, Eatchel, &Wambaugh, 2015).

Single-subject designs comprise the majority of the SPT investigations thus far with fairly consistent outcomes for individuals with a wide range of AOS severity levels (Bailey et al., 2015; Wambaugh & Mauszycki, 2010). Improved acquisition of target sounds in treated and untreated words has been observed (Bailey et al., 2015; Wambaugh & Mauszycki, 2010; Wambaugh & Nessler, 2004; Wambaugh, Nessler, Cameron, & Mauszycki, 2013). Generally, improvement noted in trained words tends to coincide with improvements in untrained words with the same segmental target, suggesting some evidence of motor learning (i.e. transfer effects from trained to untrained contexts), although a high degree of variability has been reported (Wambaugh, Martinez, McNeil, & Rogers, 1999). Transfer effects to untreated speech sounds tend to show negligible to no change (Wambaugh & Mauszycki, 2010; Wambaugh & Nessler, 2004; Wambaugh Kalinyak-Fliszar, West & Doyle, 1998). In an attempt to induce system-wide change, by improving a sound class based on manner category, limited transfer effects were also found in sentence level productions (Wambaugh, West & Doyle, 1998). Wambaugh and colleagues (1998) trained a limited number of sentence exemplars and examined generalization to untrained sentences for an adult with moderate apraxia and non-fluent aphasia in a single subject multiple baseline study. Generalization occurred to sentences containing a predominance

of the trained sound types but transfer of the trained sound types to sentences with mixed consonants was negligible (Wambaugh et al., 1998).

Although preliminary studies of SPT have been promising, seeming lack of transfer effects between related target sounds suggests a need to alter or improve on this treatment approach. Amending an SPT hierarchy to include intersystemic facilitation, and more specifically visual biofeedback, is one way that could enhance the effects of treatment. Next, we review of some visual biofeedback devices that have been used in speech treatments for AOS and the evidence for their use from clinical research.

1.5 Visual Biofeedback in AOS Treatment

Visual biofeedback is a type of augmented feedback, defined as an externally provided feedback that is separate from the intrinsic feedback that results from the movement itself (Schmidt & Lee, 1999). The motor learning literature separates augmented feedback into two categories, knowledge of results and knowledge of performance. Knowledge of results (KR) is simply feedback as to whether the intended movement was correct or incorrect. Knowledge of performance (KP) is information about the qualitative aspects of a movement (e.g., "you did not get the back of your tongue high enough"). Since the opportunity to observe the specific articulations of speech is rare, visual biofeedback provides a type of KP that is often unavailable to speakers.

AOS researchers have speculated on the relative benefits of treatment approaches that use some kind of augmented input to improve upon treatment outcomes (Katz, McNeil, & Garst, 2010; Wambaugh & Mauszycki, 2010). From a schema theoretical perspective, Katz and colleagues (2010) hypothesized that augmented feedback of speech movement could increase

motor learning by strengthening GMPs, parameters or both. Using a DIVA approach, Wambaugh and Mauszycki (2010) make reference to the model's feedback systems and, in their experiment with a speaker with severe AOS, suggest that treatment might have been more effective if it had included specific training to recognize and self-monitor somatosensory and auditory cues.

Indeed, researchers have studied various augmented visual biofeedback tools to find effective treatments for AOS. Howard and Varley (1995) reported their observations of attempting to improve tongue-palate contact in a speaker with severe apraxia using electropalatography (EPG). EPG is a type of articulatory feedback technology that uses an artificial palate embedded with electrodes that is placed in the upper dental arch during treatment so that the learner can see and modify their patterns of contact between tongue and palate. Observations of training with a single participant suggested that the speaker was able to make use of the visually provided information in order to modify their articulations of stop consonants that required specific patterns of palatal contact in words (e.g., "ladder") (Howard & Varley, 1995). Though tentative, these observations appear to provide some evidence that articulatory visual feedback could be an effective component to treatment approaches for AOS.

Electromagnetic articulography (EMA) is a non-invasive visual feedback tool that has also been used in treatment for AOS (Katz et al., 2010; McNeil, Katz, Fossett, Garst, Szuminsky, Carter, & Lim, 2010). EMA uses low intensity magnetic fields to track the movement of sensors glued to various articulators (McNeil et al., 2010) For one adult with AOS and aphasia, treatment using EMA feedback improved acquisition and retention of speech movements (Katz et al., 2010). McNeil and colleagues (2010) also explored using EMA in speech treatment for two adult

participants with AOS and aphasia. The treatment paired the KP feedback provided by EMA with the clinicians' perceptually based KR feedback and found substantial generalization to similar speech movement patterns to those that were trained (i.e. generalization from trained "g" to untrained "k") (McNeil et al., 2010). They also found some transfer of training effects to dissimilar speech movement patterns however, and did not draw firm conclusions from these outcomes (McNeil et al., 2010). Researchers using EMA have suggested that the variability in outcomes may indicate that visual feedback technology may not operate on speech motor learning in a way that is uniform for all treatment targets (Katz et al., 2010).

Providing feedback in the visual modality could focus a speaker's attention on particular locations and movements of articulators. For the current study, ultrasound was the chosen feedback device. The present evidence for its use in speech treatment is outlined below.

1.6 Ultrasound (U/S) in Speech Treatment

Ultrasound is a visual feedback technology that has been used in speech treatment for a wide range of clinical populations (Adler-Bock, 2004; Bacsfalvi, 2010; Bacsfalvi & Bernhardt, 2011; Fawcett, Bacsfalvi, & Bernhardt, 2008; Preston & Leaman, 2014; Shawker & Sonies, 1984). Ultrasound uses a transducer which emits ultrasonic waves to create images on a computer of the air just above the tongue's surface. Using ultrasound, lingual movement can be observed during speech. The transducer is held directly against the skin under the chin and the ultrasonic waves pass through bodily tissues and are reflected off the air just above the tongue surface. These reflected waves return to the transducer to create an image of a moving white line which corresponds to the shape and position of the tongue. A sagittal view can be seen in 0 This orientation is used most often in treatment and research and shows the tongue from root to tip

providing information about tongue backness and relative height. Another view, coronal, allows for observation of midline grooving and tongue height (Bernhardt, Gick, Bacsfalvi, & Adler-Bock, 2005). Shawker and Sonies (1984) reported the preliminary results of using ultrasound in speech training where they postulated that the ultrasound image could help to imprint newly learned tongue positions that are then added to acoustic and tactile awareness to establish new motor plans.

A primary benefit of ultrasound imaging is that it captures dynamic aspects of articulatory gestures in real time. In addition, ultrasound provides knowledge of performance feedback in a unique way in speech sound treatment since a participant can observe his/her own tongue movements and compare them to a clinician's tongue movements. Ultrasound is minimally invasive and requires no individual hardware, unlike EPG. The displays are relatively easy to understand and the technology is portable so it can be brought to the clinical location of convenience. In addition, it is becoming more widely available to clinicians as the technology becomes less expensive.

In treatment research the tool has been shown to be helpful especially in the remediation of English "r" in older children and adolescents (Adler-Bock, 2004; Modha, Bernhardt, Church & Bacsfalvi, 2008) and young adults with Down Syndrome (Fawcett et al., 2008). In these case studies, the ultrasound facilitated acquisition of the target sound after traditional therapy approaches had previously been unsuccessful. Ultrasound has also been used successfully in conjunction with EPG as part of a treatment program for multiple sound targets for adolescents with hearing impairment (Bacsfalvi and Bernhardt, 2011).

1.6.1 Ultrasound in AOS Treatment

Currently, ultrasound has been explored as a treatment tool for speakers with AOS in a single case report by Preston and Leaman (2014). The current study draws support for its research objectives in part based on the outcomes from this case study.

The Preston and Leaman (2014) study focuses exclusively on using ultrasound in treatment for the English "r" with a 59-year-old female participant with moderate AOS and Broca's aphasia as the result of a CVA. The participant began the study 14.5- months post onset of stroke at which time she had already received 10 months of typical speech and language treatment. At the onset of the study "r" was the only sound that she could not acquire in any context. She was reported to have strong language comprehension skills and the ability to follow directions but her speech intelligibility remained moderately impaired. She spoke mostly in monosyllabic words with sound substitutions and errors.

It was hypothesized that treatment with ultrasound would help the participant recognize the tongue configuration for an "r" (Preston & Leaman, 2014). The production of that configuration in sequences of other sounds and more complex utterances would indicate evidence of motor learning in terms of transfer of learning from the trained sound to an untrained context. In the treatment procedures reported in this study, pre-practice was provided prior to the treatment cycle. Anatomical landmarks were identified and various aspects of tongue movements required for "r" were practiced in isolation (Preston & Leaman, 2014). Four variants of "r" were explored, prevocalic /ro/, postvocalic /or/, prevocalic /re/ and postvocalic /ar/. Each was targeted in a complexity hierarchy (i.e syllabic, monosyllablic, multisyllabic and phrase level) (Preston & Leaman, 2014). Each stage of the hierarchy was attempted only after mastery (i.e., five out of six
correct productions at that level) was observed in the stage before it. In each treatment session, four blocks, A, B, C, and D of ten minutes each were completed with ultrasound being included in treatment for block A and C (Preston & Leaman, 2014).

Probe words were elicited at the beginning of every other treatment session and at least once per week for four weeks following treatment. Probe words included a list of words with rhotics, 24 of which were treated and 36 of which were untreated, and a list of control words including word-initial and word-final stop consonants.

The correct articulations of both postvocalic contexts for "r" were observed in assessment probes after 9 of 12 treatment sessions and continued to increase in three follow-up probe sessions that took place after the cessation of treatment (Preston & Leaman, 2014). Prevocalic "r" was not observed to improve in any of the probe word elicitation sessions. The probes with word-initial and -final stop consonants showed variability across probe sessions but no trend of improvement which suggests that the improvements observed for word final "r" were a result of treatment.

Although this study is limited in its scope and design, following the treatment of a single participant with AOS for a single sound target, it appears to provide sufficient impetus to continue with inquiries into the use of ultrasound in treatment with this clinical population.

1.7 The Current Study and Research Questions

The current study draws on the theories of AOS as a deficit in speech motor programming and attempts to use models of motor learning to treat acquired AOS using the articulatory-kinematic approach, SPT. The predictions of the study are based on a theoretical motor learning approach that combines aspects of a schema model of motor execution and DIVA

sensory feedback systems. It further draws on speculations into the extent of multi-sensory speech processing and feedback control in motor learning by examining the added effects that may be observed with the addition of ultrasound as an intersystemic facilitation to an SPT treatment of AOS. The major question for this study was whether ultrasound visual feedback used in conjunction with SPT for AOS can improve speech production accuracy for both treated and related targets.

SPT has been shown to result in the acquisition of multiple treated speech sound targets (Bailey et al., 2014; Wambaugh et al., 1994; Wambaugh et al., 1999). However, SPT has shown limited transfer of learning to untrained speech sound targets (Wambaugh et al., 2013). In treating one speech sound in a participant with AOS, use of ultrasound resulted in the improvements of the speech sound in one word-position, and some transfer of learning for the treated speech sound in untreated contexts (Preston & Leaman, 2014). However, transfer of learning to related motor movements is thought to be stronger evidence of motor learning than mere performance acquisition (Schmidt & Lee, 1999). In speech production, transfer of learning could potentially result between speech sounds that share articulatory/phonological features (Maas et al., 2008). The question for the current study was whether the inclusion of ultrasound in SPT training would promote generalization to untrained but related speech sound targets.

The specific research questions for the study are outlined below.

- 1. Does SPT alone for multiple treated targets have positive treatment effects for:
- a) treated speech sound targets?
- b) untreated speech sound targets that are maximally related phonologically to treatment targets?

- c) untreated speech sound targets that share some articulatory/phonological features with treatment targets that are visible on ultrasound?
- d) untreated speech sound targets that are minimally related phonologically to treatment targets and are not visible on ultrasound?

Predictions: Because SPT has shown positive treatment effects for trained speech sounds in previous treatment studies (Bailey et al., 2014; Wambaugh et al., 1994; Wambaugh et al., 1999) SPT alone was predicted to have positive treatment effects for multiple speech sound targets. However, since transfer of learning to untreated speech sound targets has been minimal in previous SPT studies (Wambaugh & Mauszycki, 2010; Wambaugh & Nessler, 2004; Wambaugh et al., 1998) we predicted that SPT alone would have positive treatment effects only for untreated speech sounds that are maximally phonologically related to treatment targets (differing in one feature only). Minimal or no treatment effects were expected for untreated targets that share few or no articulatory/phonological features, both those visible on ultrasound and those that are not. Although ultrasound was not used in this condition, the third and fourth questions allowed for comparison of outcomes for the two treatment conditions.

- 2. Does SPT plus U/S for multiple treated targets have positive treatment effects for:
- a) treated speech sound targets?
- b) untreated speech sound targets that are maximally related phonologically to treatment targets?
- c) untreated speech sound targets that share some articulatory/phonological features with treatment targets that are visible on ultrasound?

 d) untreated speech sound targets that are minimally related phonologically to treatment targets and are not visible on ultrasound?

Predictions: SPT alone has shown positive treatment effects for multiple trained speech sounds (Bailey et al., 2014; Wambaugh et al., 1994; Wambaugh et al., 1999) and ultrasound used to treat a single speech sound for an individual with AOS has shown positive treatment effects (Preston & Leaman, 2014). Thus SPT plus U/S was predicted to have positive treatment effects for treated speech sound targets. Because ultrasound allows for visual observation of lingual movement in speech and visual feedback can positively impact motor learning (Schmidt & Lee, 1999) it was further predicted that SPT plus U/S would have positive treatment effects on (1) untreated maximally phonologically related sounds *and* (2) untreated targets that share articulatory/phonological features that are visually salient on ultrasound. No positive treatment effects were expected for untreated targets that are minimally related phonologically to trained speech sound targets and furthermore, are not visible on ultrasound.

- Is there a measurable difference in the treatment outcomes between SPT alone and SPT plus U/S as evaluated by:
- a) trained listeners on an auditory perception task involving judgments of consonant accuracy?
- b) untrained listeners on a word identification task?
- c) trained observers of ultrasound video?

Predictions: Because we predicted that the inclusion of ultrasound with SPT would improve motor learning and have positive treatment effects for treated and both phonologically related as well as visually salient untreated targets, we predicted that SPT plus U/S would show

greater positive treatment effects than SPT alone as evaluated by trained listeners, untrained listeners, and trained observers of ultrasound video.

Chapter 2: Methods

A single older adult with AOS completed two blocks of treatment in a quasi-experimental case study. This study employed such a research design for a variety of reasons. First, in order to answer research question three, which compares the treatment outcomes between treatment types, two blocks of treatment needed to be completed and data collected pre- and post-treatment. Rather than comparing treatment effects between subjects, providing a single participant with both treatment types helps to increase the validity of these comparisons. As mentioned above, the previous SPT treatment studies from which this experiment was loosely modeled have also typically used single subject designs. Finally, this thesis represents initial inquiries into the effectiveness of using ultrasound within SPT for an individual with AOS and these early stages of study rely on single-subject case studies (Dollaghan, 2007).

The first block of treatment, referred to as the SPT block, followed the SPT hierarchy. The second block of treatment, referred to as the SPT plus U/S block, followed the SPT plus U/S hierarchy which was designed to follow the initial SPT hierarchy as closely as possible while including clinician modeling and participant practice using the ultrasound. Data were collected for analysis from the following experimental conditions: baseline, pre-SPT, post-SPT, pre-U/S, and post-U/S. Originally, a follow-up-U/S experimental condition was proposed in order to assess long-term maintenance of treatment effects following cessation of SPT plus U/S treatment. The participant was unable to attend this final assessment session due to a medical emergency and this data point was subsequently removed from the experimental design.

2.1 Participant

The participant (P) in this case study was a 70-year-old female with AOS and non-fluent aphasia. She was recruited from a local stroke recovery group where she was invited to participate by the program manager. At the stroke recovery center, P had been receiving Speech-Language treatment for AOS and non-fluent aphasia once per week with the Speech-Language Pathologist (SLP) at the center, who was also one of the committee members for this thesis.

At the time of her stroke, 11 months prior to the initial assessment, P had been working as a teacher/speaker. P reported no known hearing deficits, did not wear corrective lenses, and had no history of speech or language difficulties before her stroke. She was right-hand dominant but presented with right hemiparesis and used her left hand for the majority of her needs. P was highly motivated to improve her speech and participated enthusiastically in all assessment and treatment activities. P's husband acted as a caregiver. He provided support throughout the treatment by providing transportation to sessions, sharing information about P's speech and language gains and acted as the primary e-mail contact for this author throughout the study.

2.2 Initial Assessment

A speech-language assessment was administered by this author over the course of two one-hour sessions prior to the initiation of treatment. This was completed to confirm the diagnosis of AOS as well as to choose treatment targets for the SPT block of treatment. Assessment tools included the Western Aphasia Battery-Revised (WAB-R) (Kertesz, 2006), the Apraxia Battery for Adults-2 (ABA-2) (Dabul, 2000), a Dworkin-Culatta Oral Mechanism Exam (D-COME) (Dworkin-Culatta, 1980), a qualitative observation of speech behaviors that are consistent with AOS based on Wambaugh and Shuster (2001), and an elicitation of a probe word

list used to determine treatment targets. The word list is provided in Appendix B and the elicitation probe is discussed later.

ABA-2 Subtest	Diadochokinetic Rate	Increa Word	sing Length	Limb and Oral Apraxia		Utterance Time for Polysyllabic Words	Repeated Trials	
Section		А	В	А	В			
Raw Score	5	4	2	43	38	100	0	
Severity Rating	Moderate	Mild		Mild	Mild	Severe	Severe	

 Table 2-1 Results of the Apraxia Battery for Adults- 2

The ABA-2 is a normed and standardized assessment for acquired apraxia of speech (Dabul, 2000). It was used in this initial assessment to quantify the severity of AOS and collect an inventory of articulatory characteristics that would guide individual treatment. The raw scores and resulting severity ratings from the ABA-2 are found in Table 2-1. Administration of the ABA-2 also included collecting an inventory of the speech behaviors that characterized the participant's apraxia. Those behaviors that were observed to occur most often and were relevant to later treatment decisions were the following:

- perseverative phonemic errors, e.g., after attempting to produce the word "white" the sequence [wait] was produced multiple times later in the assessment;
- (2) numerous and varied off-target attempts, e.g., "key" was produced as [θi], [kre], and [wœr];
- (3) increasing errors with increasing phonemic sequences, e.g., P could imitate the word "hard" but was unable to produce "harden" or "hardening" in imitation;

- (4) marked difficulty initiating speech, typically observed as groping or false starts prior to speech production; and
- (5) intrusion of schwa between syllables or clusters, e.g., "please" produced as [helidze].

The WAB-R was administered in order to obtain an Aphasia Quotient (AQ) score and provide an opportunity for qualitative observation of P's ability to use and understand language. P showed no difficulty in responding correctly to yes/no questions (e.g., "are the lights on in this room?"), correctly pointing to named pictures or objects (e.g., "triangle", "cup", "ear"), or following simple directions (e.g., "point to the window, then the door"). P had significant difficulties in generating her own speech. Her spontaneous speech most often comprised the word "hello" and the reduplicated syllable, "dedede" which were considered formulaic and automatic utterances. In terms of language comprehension, P had difficulty following some complex sequential instructions (e.g., "point with the pen to the book"), responding to questions that included relative terms like "before" or "larger than", and identifying her left and right. The WAB-R AQ was intended to be collected as a summary of overall aphasic deficits. At the time of the assessment, P presented with an AQ of 16.7 which is considered a severe deficit. The WAB-R score was considered cautiously by this author due to the observed difficulties in generating speech which could have impacted the results for word naming and sentence completion subtests. The co-occurrence of aphasia with AOS is well documented (McNeil et al., 2009; Ogar et al., 2005; Peach, 2005; Wambaugh & Shuster, 2001). Although expected, the presence of aphasia limits conclusions about treatment outcomes. Still, attempts were made in the treatment design to control for aphasic deficits, particularly that speech was assessed and treated through repetition instead of self-generated utterances.

The three deep tests of the D-COME that were completed were Lip Functioning, Tongue Functioning, and Motor Programming Abilities. In terms of lip functioning, difficulty was observed in posturing lips for retraction and protrusion. Range of motion appeared to be within normal limits once activity was initiated. Initiation of alternative motion and lip smacking also appeared to be difficult and groping behavior was observed. Diadochokinetic rate was slow and articulations imprecise (e.g., the syllable sequence "pataka" was produced as [madata], [madaka], and [madagra] at a rate of approximately 1 per second). Lip strength appeared normal. In terms of tongue functioning, speech alternating motion rates were reduced, with abnormal syllable timing and imprecise articulations, which was also noted in the diadochokinetic productions. Non-speech alternating motion rates (e.g., repetitive lateral plane tongue movements) appeared to be within normal limits, as were range of motion and strength. Probable motor programming deficits included false starts and groping behaviors, and imprecise articulations of multisyllabic words (e.g., "motorcycle" was produced as [how.o.de.de]) and phrases (e.g., P was unable to repeat the phrase "no ifs ands or buts"). These deficits were less severe when she was provided with multiple models of a word to imitate or when given a phonetic cue (i.e., the first sounds of the word).

During standardized tests and elicitation of baseline probes, P displayed the following characteristics consistent with a diagnosis of AOS (Wambaugh & Shuster, 2001): slow speech rate, sound distortions (e.g., "window" was produced as [wɪnʤ3]), and distorted sound substitutions (e.g., "bed" was produced as [meiʤ]), as well as prosodic disturbances (e.g., inaccurate stress in multisyllabic words). In addition, she showed articulatory groping, speech-initiation difficulties, apparent awareness of errors, and repeated attempts at productions.

Standardized test and naturalistic observations of P's speech characteristics were consistent with a diagnosis of apraxia of speech and aphasia.

2.2.1 Elicitation Probe

During the initial assessment, 30 probe syllables were elicited in addition to the tests and observations above. This probe syllable list, provided in Appendix B, was also elicited before and after each block of treatment. Elicitation of these probe words in each of the experimental conditions provided the data that were used to measure treatment outcomes. The 30 syllables tested 15 different speech sound targets in word-initial position each in high and low vowel contexts. These speech sound targets were as follows: /k/, /g/, /s/, /z/, /t/, /d/, /n/, /l/, /f/, /tf/, $\frac{dz}{dy}$, $\frac{\theta}{m}$, $\frac{h}{m}$. Thirteen of the probe syllable target sounds were chosen in order to be visible on ultrasound in a mid-sagittal view (Bernhardt, Gick, Bacsfalvi, & Adler-Bock, 2005), i.e., /k/, $\left|\frac{g}{s}, \frac{z}{t}, \frac{t}{d}, \frac{h}{d}, \frac{h}{d}, \frac{h}{d}, \frac{h}{d}, \frac{h}{d}, \frac{h}{d}, \frac{h}{d}\right|$ Two additional sounds, (i.e. /m/ and /h/), were included that are not visible on ultrasound in order to be used as controls. We predicted that these two target sounds would show no improvement following treatment of lingual consonants. In order to examine transfer of treatment gains in both blocks of treatment, voiced and voiceless cognates (e.g., /k/ and /g/) were included where possible. Monosyllabic probes were used, either with open syllables, or non-lingual codas (e.g., /m/ in "theme") in order to avoid sequencing confounds (Tourville & Guenther, 2011) and make it easier to observe and interpret the ultrasound video during analysis. There is a precedent in the research on articulatory feedback to constrain treatment and generalization targets by ease of visualization (McNeil et al., 2010).

2.2.1.1 Probe Elicitation Procedure

Probe words were elicited in the same locations as in the treatment sessions. Audio and ultrasound video recordings were taken of each probe session. Audio data were recorded onto .wav files at 44.1kHz/16 bit using a Zoom H4n Handy digital recorder version 1.72 with a Sennheiser ew 100 G2 body microphone. Ultrasound data were recorded onto .cine files using an Interson GP 3.5MHz ultrasound transducer set to a frequency of 2.5Hz and a depth of 17 cm. The ultrasound transducer was connected to an Acer Aspire E15 Touch laptop which ran SeeMore ultrasound imaging software. The participant's head was not stabilized during the probe elicitations. The author acted as examiner for all probe elicitation sessions.

The baseline data were collected three weeks before SPT treatment block. The data for the pre-SPT condition were collected 7 days before the initiation of SPT treatment. The data for the post-SPT condition were collected 6 days after the final SPT treatment. The data for the pre-U/S condition were collected 28 days before the initiation of SPT plus U/S treatment. The data for the post-U/S condition were collected 3 days after the final SPT plus U/S treatment. The assessment schedule for the SPT plus U/S block is discussed in more detail in the section on SPT plus U/S treatment.

The examiner provided the written target word and asked P to repeat the monosyllables. P did not observe the ultrasound image during probe word elicitation. The examiners noted during the initial assessment that P tended to have difficulty correctly repeating words on her first attempt. Therefore, during probe elicitations, the examiner provided five chances to correctly produce the entire target syllable following the examiner's model. If the target word was not correctly produced after five attempts, the next word was provided. No explicit feedback was

provided about the target production. Implicit feedback, however, was provided, in that as soon as a target syllable was judged by the examiner to be correctly produced, no more elicitation opportunities were provided. This procedure was chosen in order to reduce frustration on the part of the participant and to provide a sense of accomplishment and success even during probe sessions. It should be noted that as a result of this elicitation procedure P produced a different total number of tokens in each of the experimental and baseline conditions. The examiner determined whether the pronunciation matched the target syllable, and transcribed each elicitation attempt using the International Phonetic Alphabet (IPA). These transcriptions were later confirmed by this author/examiner using the audio recordings taken from the sessions. The accuracy of on-line transcriptions was uncertain and the final transcriptions, provided in Appendix E, were determined after listening to the recordings multiple times. Intra- and interobserver reliability was not calculated for these transcriptions at the time of this writing because the data were not used for the main analysis related to research questions.

2.3 General Treatment Design

The theoretical models of motor learning outlined in the introduction and current research evidence on AOS treatment approaches informed treatment delivery, specifically the decisions regarding practice schedule, treatment targets, and presentation of practice stimuli in treatment. Some treatment decisions were made for pragmatic reasons based on the needs of the participant and the availability of interventionists.

2.3.1 Treatment Program

For each of the two treatment blocks, treatment was administered during 45-50 minute sessions 2 days per week for 7 weeks. One of the weekly sessions was led by a Registered

Speech-Language Pathologist, and committee member for this thesis. The other weekly session was led by the author - a speech-language pathology (SLP) graduate student. One additional weekly practice session lasting 45-50 minutes was provided with a speech-language assistant (SLA) candidate who was volunteering with the stroke recovery center at the time of this study.

Distributed practice has been reported to be more effective in speech treatment than intense massed practice (Rosenbek & Jones, 2009) and in SPT distributed and massed practice appeared to shown similar treatment outcomes (Wambaugh et al., 2013). These research outcomes allowed this author to choose practice schedules that fit within the schedules of both interventionists and the participant which aligned more closely with models of distributed practice. Therapy sessions took place in quiet and private spaces at one of two center locations. For scheduling reasons, each interventionist delivered therapy in different rooms.

2.3.2 Determining Treatment Targets

Two speech sound targets were selected for each treatment block. There were several motivations for targeting two sounds for treatment. The first was, as previously outlined, a prediction that training two targets would facilitate greater system-wide treatment effects for both blocks of treatment than training a single sound. The decision was also a response to Wambaugh et al. (1999) and Wambaugh et al. (2013) who recommended targeting more than one speech sound when applying SPT so as to limit the over-generalization effects that have been noted when treating only a single speech sound. Finally, training two targets allowed for greater potential to assess transfer effects to untreated speech sound targets.

The initial assessment baseline probe was used to determine the treatment targets for the SPT block of treatment. The pre-U/S assessment probe was used to determine SPT plus US

phase treatment targets. Transcriptions of the probe elicitation sessions are in Appendix E. In order to be selected for treatment, a speech sound needed to show more than one articulatory mismatch in the target selection probe session. In probe elicitations sessions where treatment targets were chosen, P produced speech sound errors on nearly every consonant included in the probe word list with /1/ being the only speech sound that was error-free in imitation at baseline. Considering P's extremely limited spontaneous speech production and low speech sound accuracy in baseline elicitation probes there were many suitable options for treatment targets for the SPT block of treatment. Note that, although P could produce the speech sound /d/ consistently in the formulaic utterance "dedede", she could not consistently repeat the target sound /d/ in baseline probe elicitation.

Treatment targets were chosen to be maximally different in articulatory place. This decision allowed for an examination of potential transfer effects to untreated speech sounds that share articulatory place features. In addition, maximally opposing articulatory places are easier to observe and distinguish on the ultrasound in treatment. In each phase of treatment, one treatment target was coronal (tongue-tip or blade) while the other was dorsal (velar).

For the SPT block the treated targets were the anterior coronal fricative /s/ and the dorsal stop /g/. The target /s/ was chosen for the SPT block of treatment because P's production of this sound was inconsistent at the time of target sound selection. She showed four articulatory mismatches during the baseline elicitation probe, two of which concerned articulatory place. This target sound is also common in English and was considered a functional target for treatment. The target /g/ was chosen as a maximally opposing target to /s/ because P was unable to correctly

produce that target in the baseline elicitation probe. She showed place, manner, and/or voicing mismatches in both vowel contexts.

For the SPT plus U/S block the treated speech sound targets were the coronal nasal stop /n/ and the dorsal stop /g/. The target /n/ was chosen for treatment in this block because P was inconsistent in producing this sound during the pre-U/S elicitation probe and showed both place and manner errors in both vowel contexts. In addition, P had displayed difficulty in articulating /n/ throughout the initial block of SPT treatment and expressed an interest in targeting this sound in the SPT plus U/S block. At the time of target sound selection for the SPT plus U/S block /g/ continued to show articulatory mismatches, especially in a high vowel context. In order to continue to target maximally opposing place while also leaving an untreated target to assess transfer effects, it was determined that /g/ would be selected for both treatment blocks. This also allowed for comparisons between a speech sound targeted in two blocks of treatment and a speech sound targeted in a single block of SPT plus U/S. Note that in the SPT block one stop and one fricative consonant were targeted while in the SPT plus U/S block both treated targets were stops (one oral and one nasal).

2.3.3 Presentation of Stimuli in Treatment

Speech sound targets were presented during treatment in a semi-blocked structure. The two speech sounds targeted in a single treatment session were presented in an ABA format or BAB sequence alternating for each subsequent session. In the ABA sequence, speech sound A was targeted for 11 minutes of at the beginning and end of the session, while sound B was targeted for 22 minutes in the middle of the session. The reverse pattern occurred for BAB sessions. This semi-blocked presentation was chosen both to minimize the perseveration

potential and because Knock et al. (2000) found that blocked practice did not lead to faster speech sound acquisition. Presentation of stimulus may be highly variable based on individual needs and so our chosen presentation was not expected to be detrimental to the effects of treatment.

Target words for treatment were one- and two-syllable words with the target speech sound occurring in various word positions. The target words increased in complexity as treatment progressed, with target words being removed from the treatment lists and new words added as mastery of words was achieved. Mastery was defined as correct whole word production for at least 80% of elicitation attempts in three consecutive treatment sessions. Adding new words was a way to increase novelty to the treatment sessions and reduce potential for boredom. In addition, continuing to add new treatment words with target sounds occurring in more than one word-position increased the variability of the practice contexts. According to a motor learning perspective, variability should provide opportunities for various parameters to be applied to a single GMP, thus strengthening a particular movement's schema (Maas et al., 2008). Variable practice, rather than exact practice, was expected to lead to transfer effects from training.

2.3.4 SPT Treatment Block

P attended 13 out of 14 scheduled SPT treatment sessions. The SLP chose to forgo treatment on one day due to the participant's fatigue at the beginning of the session. The participant attended seven practice sessions with the SLA candidate. Thus, 20 total treatment sessions using SPT took place during the SPT block of treatment.

2.3.4.1 SPT Hierarchy

This author developed the SPT treatment hierarchy used in treatment and practice sessions specifically for this study. The hierarchy was designed so that the same general procedure and wording used for elicitation could be replicated when using ultrasound in the following block of treatment. The SPT hierarchy can be found in Appendix C.1.

The treatment hierarchy attempted to include the techniques that are consistently reported to be components of SPT treatment studies (Bailey et al., 2015). Each step of the hierarchy was applied only if the target word was produced incorrectly in the previous step and included the following two components: modeling by the interventionist with a request for repetition before each elicitation, and oral feedback after the elicitation. Treatment hierarchy elements were applied sequentially as follows: (1) The interventionist modeled the target word in a carrier phrase and requested the participant repeat only the last word in the phrase (cloze technique); (2) the interventionist presented an orthographic cue, the target word written in black marker on a note card, and named or visually drew attention to any mispronounced speech sound; (3) the interventionist provided integral stimulation, defined by the phrase "watch, listen and say it with me" and at least one attempt by the interventionist offered articulatory placement cues and again attempted to produce the word in choral production with the participant; (4) the interventionist offered articulatory placement.

There are two additional commonly included elements of SPT hierarchies, one of which was included and the other of which was excluded from the treatment hierarchies in this study. The included component was a final step, whenever P correctly articulated a target word. When a word was correctly produced, the interventionist requested five independent productions of that

word. These five elicitations were included so that each elicited word would have a greater number of trials in which to calculate mastery. A component that was excluded from the hierarchy was minimal pair contrast practice. This author determined that the severity level of AOS would make productions of minimal pair contrasts difficult to achieve, due to interference between the two stimuli. However, articulatory placement cues did sometimes include contrasting the target speech sound with another sound produced at a different place of articulation. Therefore, minimal pair contrasting was used as a type of speech sound production cue at times, although it was not included as a specific step in the SPT hierarchy.

2.3.5 SPT Plus U/S Treatment Block

P attended 14 out of fourteen scheduled SPT plus U/S treatment sessions and six additional training sessions that included the ultrasound. During this treatment block the participant attended nine practice sessions that followed the original SPT hierarchy and included no ultrasound component. Practice sessions took place with the SLA candidate who was not trained to use the ultrasound in treatment. SPT practice sessions continued to take place during training sessions, which is why there were two additional practice sessions compared with the SPT block of treatment. A total of 20 sessions using SPT with ultrasound and nine sessions using SPT without ultrasound took place during the SPT plus U/S block of treatment.

Additional training sessions were included in this block of treatment for a variety of reasons. The motor learning literature reports that pre-practice to understand the expectations of learning can have an impact on the brain's ability to learn new motor movements (Schmidt & Lee, 1999). In the ultrasound literature, more specifically, adequate training for participants is likely to be a factor in successfully applying a treatment with ultrasound technology (Shawker

and Sonies, 1984). Two training sessions were originally scheduled, during which time P initially expressed difficulty making use of the visual feedback. Because of this difficulty, more participant training was necessary than was originally scheduled. Additionally, training sessions allowed for both clinicians to consult and confirm that the ultrasound was being applied in a consistent manner within the SPT hierarchy. This additional training and practice time is a limitation to the methodology that is considered in the discussion.

2.3.5.1 SPT Plus U/S Hierarchy

The SPT plus U/S hierarchy (Appendix C.2) was adapted from the SPT hierarchy to include ultrasound only in last two steps of the hierarchy. The therapy process for using ultrasound includes some of the same components as those found in a typical SPT hierarchy which meant that the ultrasound tool was used in a way that was as consistent as possible with the previous research on its therapeutic applications (Bernhardt et al., 2005). Specifically, the SPT plus U/S hierarchy employed: (1) direct modeling of target tongue positions and postures by interventionists, (2) participant imitation of the desired tongue shapes, and (3) freezing of ultrasound images to discuss key features of tongue shapes. The current study did not make use of palate traces affixed to the computer monitor or coronal transducer positioning, so it should be noted that not all features of ultrasound in speech treatment described by Bernhardt and colleagues (2005) were used.

Hierarchy elements that remain unchanged from the original SPT hierarchy include modeling by the interventionist with a request for repetition before each elicitation, interventionist oral feedback after each elicitation, and the first step of the elicitation protocol. SPT plus U/S hierarchy elements were applied as follows: (1) The interventionist modeled the

target word in a carrier phrase and requested repetition using the cloze technique (unchanged from SPT hierarchy); (2) the interventionist presented an orthographic cue, the target word written in black marker on a note card, plus a static articulatory visual cue, i.e., a clay model of the tongue shaped into a correct articulatory posture, naming or visually drawing attention to any mispronounced speech sound; (3) the interventionist provided integral stimulation, using "watch, listen, and say it with me", plus modeled the target word on the ultrasound and requested repetition in choral production; (4) the interventionist offered articulatory placement cues while the participant observed her own productions on the ultrasound and attempted to correct them in choral production.

Although visual stimulus was the main feedback mechanism of interest in SPT plus U/S, there was an overall greater emphasis on multi-sensory integration during this block of treatment. In addition to focusing on the visually presented input from the ultrasound image P was also encouraged to attend to the tactile and kinesthetic input associated with target speech sounds. Contrasts between alveolar and velar targets were also pointed out on ultrasound during this treatment block (i.e., minimal pair contrasts). Pre-practice was included in each SPT plus U/S treatment session. Seven minutes at the beginning of each SPT plus U/S treatment session was devoted to orienting P to anatomical landmarks and practicing isolated lingual movements using the visual ultrasound image as feedback.

2.3.6 Treatment Fidelity

Treatment fidelity checklists are provided in Appendix D. Fidelity was measured by an SLP graduate student and research assistant. Fidelity checklists were based on the treatment hierarchies from each treatment block. Measures of fidelity were collected for an entire treatment

session from each of the two interventionists involved in treatment sessions. Fidelity measures were not collected for practice sessions. Fidelity to the treatment protocol was 85.09% during the SPT block of treatment. Fidelity to the treatment protocol was 81.96% during the SPT plus U/S block of treatment. The element of the protocol that was found to be applied least faithfully for both interventionists was oral feedback following elicitation attempts. This was often the case when the participant was attempting to rapidly produce the target word multiple times.

2.4 Outcome Measures

The outcome measures that were based upon the audio-recordings of probe word elicitation attempts included whole word accuracy and speech sound accuracy. The latter also included measures of articulatory place accuracy. The outcome measure that was based on the ultrasound video-recordings was observed articulatory place accuracy.

2.4.1 Trained Listener Sound Target Accuracy Measures

Trained listener judgments were made by three SLP graduate students including this author. Ratings were completed using computerized protocols programmed in PsychoPy software version 1.83.04 (Pierce, 2007; Pierce, 2009) so that raters were blinded to the treatment condition during the judgment task. These judgments, unlike the whole word accuracy measures, were based on the accuracy of the syllable-initial speech sound only and ratings were made for all elicitation attempts – up to five attempts based on experimental stimuli elicitation procedures. Listeners were shown the target sound written in the IPA and asked to rate it on a three-point scale. The scale was associated with the responses "yes, the target sound matched the sound shown", "no, the target sound did not match the sound shown", and "the target sound was approaching the sound shown".

The inter-rater reliability for sound accuracy based on consensus between all three trained listeners was measured separately for each condition and is listed in Table 2-2.

Table 2-2 Measures of Inter-Rater Reliability for Trained Listeners

Pre SPT	Post SPT	Pre US	Post US
91/107 = 85%	68/84 = 80.96%	100/120 = 83.3%	54/61 = 88.52%

2.4.2 Untrained Listener Whole Word Accuracy Measures

Untrained listener judgments were based on only the participant's initial attempt of each probe word in treatment condition and baseline probes. Whole word identification judgments were made by three untrained listeners. Untrained listeners were adult volunteers with no reported hearing deficits who responded to a recruitment flier. Listeners were blinded to the treatment conditions of each token and asked to transcribe orthographically the word or syllable that they heard. The majority response was used to determine the percentage of correct whole words. This task was completed using computerized protocols programmed in PsychoPy software version 1.83.04 (Pierce, 2007; Pierce, 2009).

Inter-rater reliability for whole word initial attempt accuracy was 100% for two of the three untrained listeners who orthographically transcribed presented syllables. The inter-rater reliability based on all three untrained listeners who orthographically transcribed presented syllables was 52%.

2.4.3 Articulatory Place Accuracy Measures Observed on Ultrasound

A single SLP graduate student volunteer rated the articulatory place of the utterances based on the observed ultrasound video-recordings. Probe word tokens for observation included the treated speech sound targets /g/, /s/, and /n/, and corresponding untreated cognates, /k/, /z/, and /d/. The volunteer rater observed video from the pre-SPT, post-SPT, pre-U/S and post-U/S conditions. Ultrasound video recordings from the initial assessment were not saved due to technical issues so observations of baseline articulations could not be made. Videos were provided in a randomized presentation and were replayed as many times as was requested by the rater. During training the rater was informed what the CV probe words were but did not know which word was being presented for any given token.

The author facilitated training in observation of the ultrasound video, which included watching and discussing ultrasound videos of a typical speaker producing the CV probe words and an opportunity for the rater to observe their own articulations of the words using the ultrasound. Instruction was provided for how to identify the tip, coronal and dorsal sections of the tongue in static and moving images.

The task of the volunteer rater was to then observe soundless ultrasound videos of the initial token of the CV word probes and rate them based on the perceived section of the tongue that was responsible for making the observed articulation. The rating choices were coronal, which included blade and tongue tip, or dorsal, specified as mid-tongue to root. Ratings could include both articulatory places if both appeared for a given token. However, the rater was asked to indicate if one section of the tongue made articulatory contact before the other whenever possible when designating both articulatory places.

Intra-rater reliability based on the re-examination of the first ten video tokens was 90%. However, this reliability measure does not capture the limitations in analyzing ultrasound video, something that is explored further in the discussion.

2.4.4 Perceived Articulatory Place Accuracy Measures

Measures of perceived articulatory place feature matches were based on this author's narrow transcriptions of the participant's productions. These judgments were made based on this author's perceptions of the audio-recorded data, and so blinding to experimental conditions was not possible. The transcriptions used can be found in Appendix E.

2.5 Analysis

Presented in the results section is a descriptive analyses of the measured changes in articulatory accuracy for each treatment block (SPT and SPT plus U/S). Treatment effects were determined by comparing the accuracy in the pre-treatment conditions with the post-treatment conditions for whole words and target speech sounds.

Speech sound accuracy, which was based on all elicitation attempts, is reported differently. Due to the inconsistency between conditions in terms of the total number of elicitation attempts for each speech sound, speech sound accuracy could not be evaluated using percentage of correct productions. Rather, the raw number of speech sound mismatches produced in each probe elicitation are reported in the results. Statistical analysis of speech sound accuracy could not be performed due to the small and uneven number of targets. For speech sound accuracy, treated speech sounds and untreated sounds were examined separately. Untreated speech sounds were divided based on their phonological similarity and articulatory similarity to the treated sounds and observability on ultrasound. The first category included maximally

phonologically related sounds, i.e., those that share manner and place with a treated speech sounds and are also maximally similar on the ultrasound, i.e., voiced and voiceless cognates /k/ (related to /g/) and /z/ (related to /s/), and the alveolar coronal /d/, similar in tongue configuration and coronal alveolar place to /s/ (SPT only) and /n/ (SPT plus U/S). The second category included visually related sounds: affricates /tʃ/ and /dʒ/, liquids /l/ and /ɹ/ and the interdental fricative / θ /. These speech sounds have articulatory gestures that appear somewhat similar visually to a treated speech sound on the ultrasound, either because of dento-alveolar position (the /l/ and / θ /), similar to /s/ and /n/, or a more posterior position, in between /g/, /s/ and /n/ (the affricates and / μ /). Minimally related sounds were those that had no observable lingual postures on ultrasound and were used as controls. Speech sound accuracy was further analyzed in terms of articulatory place accuracy. Whole word accuracy, which was based on initially attempted productions, reports the percentage of correct productions for each condition.

One target sound, /t/, was excluded from all audio-based analysis in all conditions and baseline due to an inability to recover audio-recordings of one of the elicited syllable tokens. In addition, the target sound /ʃ/ was not included in analysis of speech sound accuracy. This speech sound was excluded due to inconsistencies between independent raters and online judgments made during the pre-SPT probe condition. More specifically, this author's online judgment of a correct initial production was not corroborated by independent ratings and so it was determined that these raw data did not accurately reflect the number of articulatory mismatches that might have been produced if continued elicitations had been attempted.

For the analysis of observed ultrasound video-recordings, tongue contact that matched the expected target place of articulation (e.g., the syllable "see" being identified as using a coronal

articulation) was identified as a match. Tongue contact that opposed the expected target place of articulation (e.g., "caw" being identified as using a coronal place of articulation) was identified as a miss. Any tongue contact that was rated as both coronal and dorsal was identified as a near match. These articulations were considered visibly imprecise. As was the case for perceptual speech sound accuracy, treated targets were separated from untreated targets. However, in the analysis of observed place accuracy only maximally related untreated speech sounds were included and so untreated targets were not divided any further.

Chapter 3: Results

Results first address research questions with respect to the trained listener auditory perceptual judgments, then whole word accuracy by untrained listeners, and finally trained observer ultrasound image ratings. Within the first section, target types are discussed in the following order: (1) treated targets, (2) untreated targets that are maximally related phonologically, (3) untreated targets that share articulatory features with treatment targets and are visible on ultrasound, and (4) minimally related targets. Within each subsection results for both SPT alone and SPT plus U/S are described.

3.1 Outcomes as Rated by Trained Listeners on Speech Sound Accuracy

Speech sound accuracy, as rated by trained listeners, can be found in Figure 3-1, 3-2, 3-3, and 3-4. These figures show the overall number of articulatory mismatches produced for up to five elicitation attempts in baseline and treatment conditions. Therefore, a reduction in articulatory mismatches indicates greater sound accuracy. In Figure 3-1 and 3-3, a triangle marker represents dorsal consonants, a circle marker represents a coronal fricative, and a square represents a coronal stop. In Figure 3-2 a circle represents /l/, an asterisk represents / θ /, a triangle represents /t/, a diamond represents /dz/, and a square represents /I/. The minimally related sounds that do not share articulatory place with the treated sounds are represented by X (Figure 3-4). Figure 3.5 shows a summary of the overall speech sound production accuracy data.

3.1.1 SPT and SPT Plus U/S: Treated Speech Sounds

The number of articulatory mismatches for treated sounds /g/, /s/ and /n/ can be seen in Figure 3-1. Between the pre-SPT and post-SPT conditions (treatment targets were /g/ and /s/ but excluded /n/) the treated speech sound /g/ showed a reduction in articulatory mismatches.

Although mismatches were observed in the baseline condition for the treated sound /s/, it remained error-free in both SPT conditions. Articulatory mismatches for the untreated /n/ showed little change prior to being targeted for treatment.

Between the pre-and post-U/S conditions (in the U/S treatment block the treated speech sounds were /g/ and /n/ but excluded /s/), the retreated sound /g/ showed a decline in mismatches equivalent to the decline between SPT conditions. The treated sound /n/ showed a sharp decrease in the number of articulatory mismatches. The previously treated sound /s/ showed a trivial reduction in mismatches because it had a low number of mismatches in the pre-U/S condition.

3.1.2 SPT: Untreated Maximally Phonologically Related Speech Sounds

The number of articulatory mismatches between pre-SPT and post-SPT conditions for the untreated maximally phonologically related sounds /k/, /z/, and /d/ can be seen in Figure 3-3. Between pre-SPT and post-SPT conditions a reduction in articulatory mismatches was observed for the untreated sound /k/, similar in magnitude to the treated sound /g/. There was a small reduction in mismatches for /z/; however, for this sound the number of mismatches in the pre SPT condition was already low. No change was observed in overall number of mismatches for untreated /d/ which remained high.

3.1.3 SPT Plus U/S: Untreated Maximally Phonologically Related Speech Sounds

The number of articulatory mismatches between the pre-U/S and post-U/S conditions for the untreated maximally related sounds /k/, /z/, and /d/ can be seen in Figure 3-3. Between the pre-U/S and post-U/S conditions each of the maximally related sounds, untreated /d/, untreated /k/, and untreated /z/ showed decreases in the number of overall mismatches that appeared similar in magnitude to treated sounds /g/ and /n/.

3.1.4 SPT: Untreated Speech Sounds Visually Salient on Ultrasound

The number of articulatory mismatches between the pre-SPT and post-SPT conditions for the untreated sounds that are visually salient on ultrasound can be seen in Figure 3-2; these included /tʃ/, /dʒ/, /l/, /a/, and / θ /, i.e., both anterior (/l/, / θ /) and post-anterior (/tʃ/, /dʒ/, /x/) coronals. The anterior coronals share place features with the treated targets /s/ and /n/, the / θ / shares manner and voicing with /s/, and the /l/ shares sonorance with /n/. The affricates and /x/ were post-alveolar, thus, still coronal, but in between the anterior /s/ and /n/ and the dorsal /g/. Between the pre-SPT and post-SPT conditions there was an apparent reduction in mismatches for the anterior-coronal /l/. The number of mismatches for post-anterior /tʃ/ and /dʒ/ showed relatively little change and remained high for both SPT conditions. Mismatches for the coronal anterior / θ / remained unchanged between pre- and post-SPT, remaining moderately low for both conditions. No articulatory mismatches were observed for the sound /x/ in either SPT condition (before or after treatment).





Figure 3-1 Mismatches for Treated Sounds

Figure 3-3 Mismatches for Maximally Phonologically Related Sounds



Figure 3-2 Mismatches for Sounds Visually Salient on U/S



Figure 3-4 Mismatches for Minimally Related Sounds

3.1.5 SPT Plus U/S: Untreated Speech Sounds Visually Salient on Ultrasound

The number of articulatory mismatches between the pre-U/S and post-U/S conditions for the untreated sounds that are visually salient on ultrasound /tʃ/, /dʒ/, /l/ /J/ and / θ / can be seen in Figure 3-2. Between the pre-U/S and post-U/S conditions, the coronal post-anterior /dʒ/ displayed a steep reduction in mismatches. The coronals /tʃ/ and /l/ also showed a decrease in number of mismatches. The anterior coronal / θ / showed no change in the number of mismatches, which remained high for both U/S conditions. No articulatory mismatches were observed /J/ in either U/S condition which was already accurate prior to treatment.

3.1.6 SPT: Untreated Minimally Related Speech Sounds

The number of articulatory mismatches between the pre-SPT and post-SPT conditions for the untreated minimally related sounds /m/ and /h/ can be seen in Figure 3-4. Between the pre-SPT and post-SPT conditions, /m/ showed a noticeable decrease in mismatches. Mismatches for /h/ remained relatively unchanged and were already low for both pre- and post-SPT conditions.

3.1.7 SPT Plus U/S: Untreated Minimally Related Speech Sounds

The number of articulatory mismatches between the pre-U/S and post-U/S conditions for the untreated minimally related sounds /m/ and /h/ can be seen in figure 3-4. Between the pre-U/S and post-U/S conditions, the number of mismatches for /m/ showed a slight decline; however, the number of mismatches for this sound in the pre U/S condition were already low. No mismatches were observed for /h/ in either U/S condition which was already quite low prior to SPT plus U/S treatment.

3.1.6 Overall Effects of SPT and SPT Plus U/S on Speech Sound Accuracy

A summary of the total number of articulatory mismatches for each group of speech sounds by condition can be found in Table 3-1. A graph of the total number of speech sound mismatches for each condition can be seen in Figure 3-5. Between the pre-SPT and post-SPT conditions there was a clear decline in speech sound mismatches. Between the pre-U/S and post-U/S conditions there was an even steeper decline in speech sound mismatches. There was a noticeable increase in mismatches following the cessation of SPT treatment before the initiation of SPT plus U/S which can be seen between the post-SPT and the pre-U/S conditions.

Table 3-1	Summary	of Articulator	y Mismatches
-----------	---------	----------------	--------------

	Baseline	Pre-SPT	Post-SPT	Pre-U/S	Post-U/S
Treated Sounds	20	17	10	17	2
Phonologically Related Sounds	19	27	11	35	7
Visual. Related Sounds	18	25	23	34	15
Min. Related Sounds	5	7	1	2	0
Total Articulatory Mismatches	62	76	45	88	24



Figure 3-5 Total Mismatches for Treated and Untreated Sound Targets

3.2 Outcomes as Rated by Untrained Listeners on Whole Word Accuracy

Whole word accuracy of the initial elicitation attempt was rated by untrained listeners.

The data are reported in Table 3-2 for the baseline and treatment conditions.

Table 3-2 Percentage of Correct Productions for Initial Elicitation Attempt

Baseline	Pre-SPT	Post-SPT	Pre-U/S	Post-U/S
6.90%	13.79%	44.82%	20.69%	65.52%



Figure 3-6 Percentage of Correct Productions for Initial Elicitation Attempt

The graph in Figure 3-6 shows the percentage of correct initial productions. An increase is observed between the pre-SPT and post-SPT conditions as well as between the pre-U/S and post-U/S conditions. Whole word accuracy of the initial elicited attempt reached 44.83% in the post-SPT and in the post-U/S condition reached 65.52%. In both cases there was approximately a 30% increase in initial production accuracy between pre- and post- conditions; although, as the figure shows, the SPT plus U/S block started with a higher accuracy than the SPT block.

Between the post-SPT and pre-U/S conditions, following a three-week cessation of SPT treatment, there was a decline in the accuracy of initial productions of approximately 20%. This pattern of improvement followed by a decline in accuracy with withdrawal of treatment was similar to the pattern observed in speech sound accuracy.

3.3 Evaluation of Ultrasound Videos

Graphs of observed articulatory place accuracy as rated by trained observers can be seen in Figure 3-7, 3-8, 3-9, and 3-10. The observed articulatory place accuracy for the pre-SPT condition for treated and untreated speech sounds can be seen in Figure 3-7. The words with treated sounds (/g/ and /s/) were distributed along a continuum with slightly more articulatory place misses observed than matches or near matches. The words with untreated sounds (/k/, /z/, /d/, and /n/) were predominantly judged as near matches in articulatory place. The observed articulatory place accuracy for the post-SPT condition is shown in Figure 3-9. There was an increase in the number of words with treated targets that were judged to accurately match articulatory place. There were more words with untreated sounds that were found to miss the articulatory place than in the pre-SPT condition.

The observed articulatory place accuracy for the pre-U/S condition for treated and untreated speech sounds can be seen in Figure 3-8. None of the words with treated sounds (/g/ and /n/) were observed to accurately match articulatory place. Untreated sounds (/k/, /d/, /z/, and /s/) were predominantly observed to nearly match articulatory place, similar to the pre-SPT condition. The observed articulatory place accuracy for the post-U/S condition can be seen in Figure 3-10. There were more words with treated sounds that matched articulatory place than there were in the pre-U/S condition. The words with untreated targets show a similar distribution of observed place accuracy in the post-U/S condition as in the pre-U/S condition.


Figure 3-7 Pre-SPT Observed Accuracy of Place



Figure 3-8 Pre-U/S Observed Accuracy of Place



Figure 3-9 Post-SPT Observed Accuracy of Place



Figure 3-10 Post-U/S Observed Accuracy of Place

3.4 Effects of SPT and SPT Plus U/S on Accuracy of Articulatory Place

Although not proposed as one of the main research questions, overall perceived articulatory accuracy was examined further based on accuracy of articulatory place features. These graphs can be seen Figure 3-11, 3-12, 3-13, and 3-14. In Figure 3-11 and 3-13 a triangle marker represents dorsal consonants, a circle marker represents a coronal fricative, and a square represents a coronal stop. In Figure 3-12 a circle represents /l/, an asterisk represents / θ /, a triangle represents /tf/, a diamond represents /dz/, and a square represents /I/. The minimally related sounds that do not share articulatory place with the treated sounds are represented by X (Figure 3-14).

The number of articulatory place mismatches for the treated speech sounds /g/, /s/ and /n/ can be seen in Figure 3-11. The graph shows that between the pre-SPT and post-SPT conditions the dorsal /g/ showed a steep decline in the number of articulatory place mismatches. Prior to being treated, the coronal /n/ shows a slight increase in place mismatches. The treated coronal /s/ showed no place mismatches in either SPT condition despite having shown place mismatches in the baseline condition.

Between the pre-U/S and post-US conditions both treated sounds, the dorsal /g/ and coronal /n/ showed a decrease in the number of place mismatches, similar in magnitude to the decline observed between the SPT conditions for /g/. The previously treated speech sound /s/ was accurate in terms of place errors prior to treatment and continued to showed no place mismatches in either of the U/S conditions.

The number of articulatory place mismatches for the untreated maximally phonologically related sounds /k/, /z/, and /d/ can be seen in Figure 3-13. The graph shows that between the pre-SPT and post-SPT conditions the dorsal /k/ showed a reduction in place mismatches that appeared to be similar in magnitude to the reduction observed for the treated dorsal /g/. There was a slight decrease in number of place mismatches for the coronal /z/; however, this target, like /s/, had a low number of place mismatches in the pre-SPT condition. There was an apparent small increase in the number of articulatory place mismatches for the coronal /d/.

Between the pre-U/S and post-U/S conditions there was some amount of decline in the number of place mismatches for all of maximally phonologically related speech sounds. The dorsal /k/ showed a noticeable reduction in place mismatches. The coronal /d/ showed a less noticeable reduction in place mismatches. The coronal /z/ was already fairly accurate in terms of place errors in the pre-U/S condition and showed the least noticeable reduction in articulatory place mismatches.

The number of articulatory place mismatches for untreated speech sounds that are visually salient on ultrasound /tʃ/, /dʒ/, /l/ / x/ and / θ / can be seen in Figure 3-12. The graph shows that between the pre-SPT and post-SPT conditions, the anterior coronal /l/ showed a decrease in place mismatches. The post-anterior coronal /tʃ/ showed a slight decrease whereas /dʒ/ showed a slight increase in the number of place mismatches. The anterior coronal / θ / showed no change in place mismatches between SPT conditions and remained relatively low for both SPT conditions.

Between the pre-U/S and the post-U/S conditions the speech sound /l/ showed a decrease in the number of articulatory place mismatches, similar in magnitude to the decline observed between the SPT conditions for this target. The coronals /tʃ/ and /dʒ/ also showed a decline in the

number of place mismatches. The /I/ (Labial-Coronal-Dorsal) remained free of articulatory place mismatches for all conditions. The interdental / θ / showed no change in number of place mismatches which remained high for both U/S conditions. It is the only sound in this group that did not show a reduction in place errors between the pre-U/S and post-U/S conditions.

The number of articulatory place mismatches for minimally related sounds /m/ and /h/, can be seen in Figure 3-14. The graph shows minimally related speech sounds were consistently low in all conditions. Between the pre-SPT and post-SPT conditions the bilabial /m/ showed a reduction in place mismatches. The glottal /h/ showed a slight decrease in mismatches; however, mismatches for this sound was already very low in the pre-SPT condition.

Between the pre-U/S and post-U/S conditions the bilabial /m/ showed a slight decline in articulatory place mismatches. The glottal /h/ remained free of articulatory place mismatches for both U/S conditions.



Figure 3-11 Place Mismatches for Treated Sounds







Figure 3-12 Place Mismatches for Sounds Visually Salient on Ultrasound

Sounds



Figure 3-14 Place Mismatches for Minimally Related Sounds

Chapter 4: Discussion

This case study explored the effects of SPT treatment with and without ultrasound. It was expected that the use of a visual biofeedback device as intersystemic facilitation in an articulatory-kinematic treatment could improve speech sound accuracy for an adult with AOS. Treated and untreated speech sounds were assessed to infer motor learning in terms of transfer effects from trained speech sound targets to untrained speech sound targets. The following section discusses the reported results of the study. Treated speech sounds and untreated speech sounds are examined separately with respect to our initial research questions which is followed by a discussion of whole word accuracy as rated by untrained listeners. Additional outcomes of the study will be presented including measures of articulatory place accuracy as perceived from audio-recording and observed on ultrasound video. Overgeneralization and maintenance effects of SPT will be discussed. This author's qualitative observations of treatment will also be presented, including a discussion of the proposed benefits and challenges of using ultrasound in treating AOS and more specifically within an SPT hierarchy. Finally, the limitations are presented followed by overall conclusions and clinical implications of the study.

4.1 Outcomes of Treatment as Related to Research Questions

The research questions for this study were whether SPT alone and SPT with ultrasound had positive treatment effects for multiple treated targets, untreated maximally phonologically related targets, untreated targets that share articulatory features that are visually salient on ultrasound, and/or untreated targets that are minimally phonologically related and not visible on ultrasound. The additional question that was posed was whether there was a measurable difference in outcomes between SPT alone and SPT with ultrasound for any of these speech

sound targets. The following section discusses the analyses of treatment outcomes with respect to perceived speech sound accuracy and observed articulatory place accuracy. It should be noted that there were no controls in place to accommodate for the cumulative effects of the previous SPT treatment block on the measurements of the SPT plus U/S block of treatment. For this reason, all of the outcomes of this study should be interpreted very cautiously.

4.1.1 Treated Speech Sounds

It was hypothesized that SPT used with and without ultrasound could induce motor learning for two treated speech sounds as a result of treatment. It was suggested that this learning would lead to more accurate articulation of treated speech sounds. This would be observed as an increase in the accuracy of articulation for treated speech sound in the post-U/S condition compared to the pre-U/S condition and in the post-SPT condition compared to the pre-SPT condition. It was predicted that the observed effects would be equivalent to one another.

The outcomes appeared to show that treated speech sound targets showed acquisition effects in both the SPT and the SPT plus U/S blocks of treatment evidenced by a decrease in perceived acoustic errors as rated by trained listeners. The speech sounds targeted in the SPT treatment block, /g/ and /s/, showed a decline in mismatches after a single block of SPT treatment, which was then followed by an increase in speech sound mismatches coinciding with the withdrawal of treatment, suggesting that improved speech sound accuracy was the result of SPT, although maintenance was not obtained. Although there were mismatches observed at baseline, the near lack of articulatory errors for /s/ following the initial assessment makes it difficult to draw conclusions about the effects of treatment for the initial SPT treatment block on the treated anterior coronal. Following the SPT plus U/S block, both treated targets, /g/ and /n/,

showed improvements in articulatory accuracy. The speech sound, /g/, had been treated in the previous block with SPT but improvements had not been maintained. Following the SPT plus U/S block of treatment, gains in accuracy again occurred. The speech sound, /n/, also appeared to show treatment effects after being targeted for a single block of SPT plus U/S treatment. The argument that these results are the consequence of the cumulative effects of treatment can be made and this limitation to the overall methodology of the study will be discussed further. However, the outcomes appeared to be positive following a block of SPT plus U/S for both a previously treated but unmaintained sound /g/, and a previously untreated and mismatched speech sound /n/.

It was predicted that P would acquire both treated speech sounds as a result of SPT applied with and without ultrasound. Traditional SPT studies have shown robust acquisition effects for treated speech sound targets (Bailey, Eatchel, & Wambaugh, 2015) and the results from the SPT block of treatment appeared to replicate these previous findings within the treatment condition itself. The perceived reduction in articulatory mismatches for treated speech sounds using SPT with and without ultrasound appeared to be about equivalent. SPT alone may be as effective as using STP with ultrasound in improving the articulatory accuracy of treated speech sounds, at least in the short term within treatment conditions.

4.1.2 Untreated Speech Sounds

It was hypothesized that SPT used with and without ultrasound could lead to a certain amount of positive treatment effects for untreated speech sounds that were similar to the treated speech sounds. Positive treatment effects were observed as improved articulatory accuracy in the post-treatment conditions compared with the pre-treatment conditions for untreated speech

sounds that shared the following with treated sounds: phonological features,

articulatory/phonological features visible on ultrasound, or both.Specifically, it was predicted that SPT would show positive treatment effects for untreated speech sounds that were maximally phonologically related to treated targets but that no positive treatment effects would be seen for speech sounds that were minimally phonologically related to treated targets or for speech sounds that shared articulatory/phonological features with the target, and that could be viewed on ultrasound. It was also predicted that SPT plus U/S would show positive treatment effects for both untreated sounds that were maximally phonologically related to treated speech sounds and those that shared articulatory/phonological features that could were visible on ultrasound but not for the minimally related sounds. Comparatively, it was predicted that SPT plus U/S would show transfer of learning to a greater number of untreated speech sounds than SPT alone based on trained listener ratings. This exploration of transfer effects was intended to provide further evidence of motor learning as a result of treatment.

In the SPT treatment block, treatment effects appeared to transfer to some extent from the treated speech sounds (i.e., /s/ and /g/) to the maximally phonologically related speech sounds, (i.e., /z/ and /k/). There can be few conclusions made regarding treatment effects to either the voiced or voiceless coronal fricatives during the SPT block of treatment because virtually no change in the number of errors for either speech sound was seen in this treatment block. The accuracy of these fricative was already high. The dorsal stops appeared to show more obvious transfer effects from the initial block of SPT. Mismatches for the untreated speech sound /k/ appeared to reduce sharply at the end of the block of SPT that was targeting its voiced cognate /g/. Except for minimally related /m/, the maximally phonologically related sounds, however, are

the only ones that showed transfer of learning effects in the SPT block. Speech sounds with articulatory features that were similar to the treated sounds and visible on ultrasound, (i.e., two affricates, two liquids and one inter-dental), were observed to show little to no change in articulatory accuracy between the pre- and post-SPT conditions. The reported treatment gains that were made to the treated speech sound targets during SPT did not appear to impact those speech sound targets that shared articulatory gestures with the treated targets. For the speech sounds that were minimally phonologically related to treated targets and not visible on ultrasound, /m/ did improve notably. This was either a transfer of learning effect during the SPT block of treatment or spontaneous change unrelated to treatment. With the exception of /m/, the findings of transfer from treated speech sounds to the maximally phonologically related speech sounds to the study.

In the SPT plus U/S block, transfer effects to related speech sounds were more discernible than in the block of SPT alone. Transfer appeared to extended to both maximally phonologically related speech sounds as well as to the speech sounds that shared articulatory features that are visible on ultrasound. There was a decline in errors for the untreated /k/ during the SPT plus U/S treatment that also targeted /g/. This relationship was a similar pattern to the one observed for these sounds for the SPT block. In the SPT plus U/S block, while the coronal sound target was /n/, pronounced transfer effects were observed for all untreated coronal stops and fricatives. The untreated speech sound /d/, the previously treated speech sound /s/, and its voiced counterpart /z/ all decreased in their number of articulatory errors. The declines in articulatory errors for the two coronal fricatives after the SPT plus U/S block should be considered carefully, however. The decline in errors for /s/ is minimal and the coronal fricative

/z/ may have been subject to the effects of re-acquisition. The reduction in articulatory errors could be a response to "booster" treatments which have been recommended to increase retention of treatment gains in SPT (Wambaugh & Mauszycki, 2008; Wambaugh et al., 1999).

However, evidence of transfer of learning to untrained speech sounds also appeared evident for the sounds that shared articulatory features visible on ultrasound for the SPT plus U/S block of treatment. The untreated speech sounds that composed of similar lingual articulations to the treated sounds that are also observable on ultrasound showed a trend of improvement for SPT plus U/S. More specifically, the coronal affricates and the lateral showed improved speech sound accuracy in the treatment that included ultrasound which is a trend was not observed in the treatment using SPT alone. These positive treatment effects were not observed for every speech sound that was visible on ultrasound. The interdental θ showed consistently low articulatory accuracy and the rhotic /1/ showed consistently high accuracy in both U/S conditions. The θ / is relatively infrequent in English, which may have accounted for its lack of change, even with attention paid to coronal place on ultrasound (less opportunities to attempt it in daily life). The affricates may have benefited from attention paid to both anterior /n/ and dorsal /g/, the coronal affricates being intermediate in position between those targets. The minimally related speech sound targets /m/ and /h/, that were not observable on ultrasound also showed little change in accuracy but were already quite accurate before the SPT plus U/S block of treatment. The fact that the majority of speech sounds sharing observable articulatory (coronal/dorsal) gestures did appear to improve may be an indication that transfer occurred selectively as a result of treatment. These findings might suggest that SPT that includes ultrasound may help improve the accuracy of untreated speech sounds that are sufficiently similar to treated speech sounds in terms of

phonological features and observable articulatory features. The results are generally in line with the predictions made about transfer treatment effects for SPT plus U/S treatment.

In both blocks of treatment transfer of learning appeared to extend to maximally phonologically related untreated speech sounds. Only for the treatment that included ultrasound was there evidence of transfer to speech sounds that shared articulatory features visible on ultrasound. This finding may be a modest suggestion that the ultrasound used in SPT was able to induce greater speech sound accuracy for some untreated consonants than traditional SPT applied without ultrasound. It was speculated in this thesis that the motor speech system would make use of the augmented sensory input that was provided to stimulate learning for untreated speech sounds. When the sensory information provided was acoustic only, as was the case with traditional SPT, learning appeared to transfer to phonological related speech sounds. In other words, only the phonology similarity of a sound impacted transfer effects when SPT was applied alone. When the sensory information provided was both acoustic and visual, as was the case for SPT plus U/S, transfer of learning appeared to include both phonologically related sounds as well as sounds that are visibly similar to treated sounds. In other words, both phonological and visible features of articulations impacted transfer effects when SPT was applied with ultrasound.

4.1.3. Whole Word Accuracy

The goal of speech therapy is ultimately to improve functional communication outcomes in order to positively impact social participation. The effects of treatment on initially attempted whole word accuracy as rated by untrained listeners was intended as a measure of overall intelligibility. This outcome measure was chosen as a way of determining if treatment had an overall impact on functional speech outcomes. Untrained listener measures have been used in

speech research as a way of measuring enhanced communication in everyday life (Bernhardt et al. 2005a).

The general trend in the accuracy of word production was observed to be similar for SPT treatment applied with and without ultrasound. Both treatment blocks showed relatively equivalent gains between the pre- and post-treatment conditions. The added benefit of ultrasound to SPT in terms of overall speech intelligibility may be negligible.

4.2.4 Place of Articulation

This section deals with the research question of treatment effects as observed on ultrasound video by a trained listener. Treatment effects, in this case, were defined by observed accuracy of articulatory place and in this section the results are discussed in relation to the perceived articulatory place accuracy as judged by the author in narrow transcriptions of audio recordings. These transcriptions can be found in Appendix E. The examination of place of articulation accuracy was, in part, in response to the reports that place of articulation is the most common error type in AOS (Ogar et al., 2005).

It was observed that articulatory place errors were the only type of error that increased noticeably in either block of treatment and this increase occurred between the pre-SPT and the post-SPT conditions. In contrast, articulatory place showed a trend of decreased errors for the SPT plus U/S block of treatment. The decline in place errors was more consistent for SPT plus U/S than the change in errors to articulatory place observed with SPT alone. During both treatment blocks articulatory place mismatches were generally unchanged for the minimally related untreated sounds which is similar to the results for speech sound accuracy. Improved overall speech sound accuracy was not isolated to improvements to place features. Voicing and

manner showed also showed a reduction in the number of mismatches present between pre- and post-conditions.

Visual observations of ultrasound video to explore place accuracy did not produce easily interpretable outcomes. To some extent, the ultrasound visual ratings appeared to deviate from the acoustic results. Based on the motor movements of speech as seen on the ultrasound video, observable improvements to articulatory place features were unverifiable. However, acoustically based perceptual judgments of speech sounds did appear to show improvements in articulatory place for each of the treatment blocks. It does appear from the comparison between the acoustic and ultrasound video data that speech sound improvement does not correlate in a systematic way with visible features of speech. Additional training for the volunteer rater on articulatory gestures and their observable consequences on ultrasound might have allowed for more sophisticated ultrasound video analysis. Still this lack of clear observable improvement in speech sound articulatory accuracy as rated on ultrasound video is a notable result. The converse effect, where observable changes visible on the ultrasound are not yet acoustically perceivable ("covert contrasts"), have been reported in ultrasound treatment studies in the past (Bacsfalvi, 2010). However, there no clear explanation for this inconsistency between visually observed and acoustic results.

4.2 Additional Outcomes of Treatment

As described in the methods section, the current study was unable to assess maintenance of treatment effects for both treatment blocks since we were unable to take follow up data after cessation of SPT plus U/S treatment. However, data was collected after the withdrawal of SPT treatment in the pre-U/S condition which allows for a brief discussion of maintenance of

treatment gains following SPT without ultrasound. The following section does not refer to a specific research question in this thesis.

4.2.2 Maintenance and Overgeneralization in SPT

In measures of both whole word initial attempt and individual speech sounds, withdrawal of SPT treatment was associated with a noticeable decline in rated accuracy. This decline suggests that the improvements made between the pre-SPT and post-SPT conditions were the result of the applied treatment and not due to spontaneous recovery. However, the magnitude of the decline in accuracy three weeks after the cessation of SPT also suggests that the maintenance of treatment gains from SPT alone was minimal. Although we were unable to make reliable comparisons between the two treatment blocks in terms of maintenance, the apparent loss of treatment gains following withdrawal of SPT alone appeared to align with previous SPT research on maintenance (Wambaugh et al., 1999).

A possible explanation for this lack of maintenance are the apparent effects of overgeneralization that was observed between the post-SPT and the pre-U/S (SPT maintenance) conditions which have also been described in previous SPT research (Wambaugh et al., 1999), Recall that during the SPT block of treatment, /s/ was targeted for treatment based on the number of articulatory mismatches that were observed in the baseline condition but that this sound showed no articulatory mismatches in either SPT condition while being treated. Directly following treatment, as well as three weeks after, in the pre-U/S condition, a considerable portion of the other speech sounds in probe words surfaced with a word-initial /s/ (e.g., "she" surfaced as [si], "D" surfaced as [si], and "knee" surfaced as [si]). This overgeneralization of treatment appeared to impact the results for the maintenance phase of the SPT (i.e., the pre-U/S condition)

more than in the acquisition phase of SPT (i.e., the post-SPT condition), thus showing a pattern of increased speech sound mismatches coinciding with the withdrawal of treatment. In the pre-U/S condition the speech sound /s/ was produced in error instead of another target sound in a total of 45 tokens out of 127 attempted productions. These outcomes can be seen in 00vergeneralization of treatment does not appear to be isolated to the SPT block of treatment. In the post-U/S condition, after /n/ was targeted in treatment, the sound /d/ surfaced with a wordinitial /n/ in both vowel contexts. However, the lack of data following withdrawal of SPT plus U/S treatment limits our ability to make concrete comparisons about the potential effects of ultrasound in SPT on maintenance or over-generalization.

4.3 Qualitative Impressions of Treatment

The following sections report the subjective impressions of this author as gathered from treatment notes.

4.3.1 Sound Production Treatment for AOS

Based on the treatment notes collected by this author, P showed improvements in achieving target speech sounds during treatment sessions using SPT and reached mastery criteria for 14 out of 29 treated /g/ words and 19 of 36 treated /s/ words by the end of the SPT treatment block. The basic SPT treatment protocol provided a consistent structure to treatment that was able to be applied consistently between interventionists with varying levels of clinical experience inferred by the high fidelity rating that was calculated for this treatment block. The drill-based model of treatment might be appropriate for training highly motivating or particularly functional words (e.g., important names or places).

A skill in which the participant showed unanticipated gains during treatment was in writing. This outcome has not been reported or investigated in the other SPT studies that were reviewed. It was unclear if this was in any way a consequence of treatment. It had been documented in this author's treatment notes in the fourth week of SPT sessions that P mistakenly spelled the target word aloud when the initial repetition was requested in the cloze phrase. This behavior occurred for the first two stimuli presented in that treatment session and then did not reoccur during that or subsequent treatment sessions. This author's treatment notes documented initial reports from the participant's husband of an improved ability to use writing on the first day of U/S treatment following ultrasound training sessions. Following observations of this development, writing was used as an avenue of communication in treatment sessions with this author to ask and answer questions. A thorough examination of P's reading and writing ability was not undertaken at the start of this study because it was beyond the intended scope of this thesis. Unfortunately, no conclusions can currently be made about the causes or implications of this observation.

4.3.2 Ultrasound in Treatment for AOS

Based on the treatment notes collected by this author, P showed improvements in achieving target speech sounds during treatment sessions using SPT with ultrasound and reached mastery criteria for 18 of 33 treated /g/ words and 23 of 34 treated /n/ words by the end of the SPT plus U/S treatment. Observations and reports from treatment indicated that P showed slow progress in becoming acclimatized to using the ultrasound in treatment. By the end of the treatment block with ultrasound, the tool was observed to be an engaging component of treatment. It provided discrete goals to practice (e.g., "bunch the tongue up") and offered novelty

to the drill-based practice. In the final treatment sessions P was able to recognize features of her own tongue movement and could self-cue using collateral hand motions (i.e., she used her left arm to try to mimic the back of the tongue raising). For P the ultrasound in treatment appeared to be particularly useful when she was able to include various other intersystemic modalities in concert with the visual image. Particularly when P integrated collateral cueing in the form of gross motor movements of her body or left arm, she seemed to have greater success in controlling her lingual movements. The addition of 3-D clay modeling in the SPT plus U/S hierarchy also appeared to be of value in helping the participant better understand how the movement of the visual ultrasound image relates to the actual articulatory gestures of the tongue.

Using ultrasound in treatment for AOS had some drawbacks as well. It was noted in the methods section that there was a training period prior to the onset of treatment that was lengthier than was originally scheduled. This was in part due to difficulty that the participant had in interpreting the image at first. P's initial difficulties in understanding how to be make use of the tool may be indicative of a potential challenge in using ultrasound for speech therapy for adults with AOS and aphasia. Explaining and interpreting the ultrasound image may require more language comprehension capacity than a traditional drill-based practice. For those with deficits from aphasia, this may result in increased cognitive load that precludes the ultrasound from being utilized to the greatest effect in treatment. With training, P became more confident interpreting and attempting to use the ultrasound image. However, this training was more time-consuming than expected.

Technical difficulties that occasionally prolonged treatment or assessment were a pragmatic challenge that are an undeniable part of using technological devices in speech

treatment. Additional accommodations that had to be made when using ultrasound in the SPT hierarchy included organizing seating to reduce glare on the computer screen, and coordinating the transducer between the clinician and the participant. Recall that due to right side hemiparesis the participant did not hold the tool herself during treatments or assessments.

4.3.2.1 Ultrasound in Sound Production Treatment

Including ultrasound in the SPT hierarchical design generally worked well. By designing the ultrasound components of treatment to be included only in later stages of the SPT plus U/S hierarchy, it was assured that visual feedback would not be applied for every token, which has been found to be detrimental to motor learning (Ballard & Robin, 2007). In addition, familiarity with the drill-procedure of SPT might allowed for consistency between early and later stages of the hierarchy.

Initial difficulty with knowing how to best apply ultrasound in the SPT hierarchy was evident on the part of the clinician, this author. Similar to the participant, this author learned to improve cueing techniques and make use of the ultrasound technology in better ways as treatment continued. Training from a clinician skilled in using ultrasound was of significant benefit to this author in applying ultrasound in a speech treatment.

There were some challenges around being faithful to an SPT hierarchy while using ultrasound. This was reflected in the measures of treatment fidelity where the U/S phase showed lower levels of fidelity to protocols. Reductions in fidelity measures during the SPT plus U/S block of treatment were often associated with a lack of interventionist verbal feedback. This was noted on occasions when the participant was producing target words very quickly. At times,

however, verbal feedback was limited intentionally allowing P to do more self-cueing and selfcorrecting.

4.4 Limitations of the Study

This report is a case study of a single participant and as such it has certain inherent limitations to external validity. The variability of the raw reported data for this participant also reduces the validity. It would be unwise to generalize these results to the population of people with AOS. Further inquiry with far greater methodological rigor would be recommended.

A primary methodological limitation to this study was the treatment block design. Due to the fact that the SPT with ultrasound was applied after the participant had already received a full block of SPT, the conclusions drawn from treatment in the SPT plus U/S block of treatment remain in question. Although attempts were made to control for this methodological restriction by treating one different speech sounds in each condition, it is necessary to consider that this limitation influenced the outcomes of the study.

A second limitation to methodology was the inability to assess maintenance of treatment effects for the SPT plus U/S block of treatment. It is presumed that an examination of retention of treatment gains might have provided a better indication of whether motor learning had taken place as a result of treatment. Although a final probe elicitation session was scheduled for three weeks following the final SPT plus U/S treatment session, the participant was unable to attend due to a medical emergency and this condition was removed from the treatment study.

Recall that probe syllables were elicited only until they were judged to be correct by the examiner in order to contribute to the participant's sense of success. There were multiple limitations to this procedure that became clear after initial assessments were undertaken and may

have impacted the outcomes of the study. The first issue with this procedure, which has been mentioned, is that it led to an inconsistent number of probe word elicitations across treatment conditions. Although most SPT studies have reported results in terms of percentage of correct elicitations (Bailey et al., 2015) this inconsistency in the number of elicitation between experimental conditions precluded this study from reporting on this measure. This is the reason that raw data on articulatory errors were reported in the results and formed the basis for the analyses. A second limitation in regard to this procedure is that during assessment sessions, correctness was determined by online judgments of word accuracy made by the examiner, this author, who was not blinded to the treatment conditions during the assessment. Although blinded ratings occurred after the assessment data was collected, there were cases where online accuracy judgements did not match with blinded rater judgements. This was the case for f/ and is the reason $\int \int dx$ was removed from analysis. The final noted limitation to this procedure was that accuracy judgments made during assessment sessions were based on whole word accuracy. An intrusive coda or other aberrant aspects of the whole word led to a greater number of elicitations even if the initial target sound was accurate.

This study used a portable ultrasound in a clinical setting and as such did not make use of specialized chairs with head stabilization systems that have been used in some other ultrasound studies (Adler-Bock, 2004). This lack of continuity between ultrasound video tokens precluded computational analysis of tongue movements or postures. For this reason, subjective observations of a volunteer rater were used as the outcome measure for ultrasound video. These observationally based outcome measurements were determined to be inconclusive; however, there remained some concern around the validity. It has been noted that the volunteer rater had

only a preliminary knowledge of ultrasound prior to being trained for this experimental task. The rater reported difficulty in recognizing when articulatory contact was being made for some of the tokens and requested each video be replayed multiple times.

Another limitation to the analysis that will be noted is that although probe elicitations and treatment hierarchies were applied in repetition as an attempt to control for any influence from deficits related to aphasia, it cannot be certain that treatment outcomes are due to an improvement in motor speech control and not an overall improvement in language processing. The observation that P's writing ability improved around the time the SPT plus U/S block began could point to a trend of improving language generation that could have affected speech production as a consequence. Although a welcome advance for P's communicative abilities, her gains in writing call into question our suggestions that treatment outcomes were due to motor learning.

The removal of speech sound data from analysis also limited the strength of the conclusions that can be drawn from this study. It has been mentioned that tokens that included the speech sound /t/ were removed from all acoustic analyses and tokens that included the speech sound / \int / were also removed for the trained listener judgments. These untreated and maximally related speech sound targets would have provided more opportunities to examine transfer of treatment effects. In addition, it has been noted that the speech sound / \int / has been found to be challenging to treat using AOS (Wambaugh et al., 1999). This is another reason it would have been of use to include / \int / in the analysis. Although, there were improvements that were observed for /tf/ and /dz/, two closely related and also difficult to treat sounds, for the SPT plus U/S treatment block.

Finally, some speech sound targets showed a low number of articulatory mismatches in the pre-treatment or baseline conditions. For these target speech sounds (e.g., /h/, /m/, /s/ and /z/)., only small improvements could be observed between pre- and post-treatment. This leads to difficulty in making comparisons of treatment outcomes with the speech sounds that showed a high number of articulatory mismatches in the pre-treatment conditions (i.e., /g/, /n/, /k/, and / θ /).

4.5 Future Directions

Future studies may be able to further explore the potential benefits of using ultrasound in an SPT treatment hierarchy or other speech treatments for AOS, especially with studies with more powerful methodological designs. Including a greater number of participants and randomizing their presentation of each treatment block could be components of a more effectively designed study. An examination of maintenance effects would also add to a future study's substance. Motor learning using ultrasound could be better assessed if both transfer and retention effects could be explored. It was unfortunate that this additional analysis could not be undertaken during this study. A study of ultrasound in clinical speech treatment would benefit from applying a stabilization technique or other process that can ease analysis of the collected video data. Stone (2005) reports some useful clinical procedures that could be used in a study that uses ultrasound in treatment outside of laboratory settings. A future study using one of these techniques might be able to deduce more valuable information from ultrasound video.

4.6 Conclusions and Clinical Implications

This case study was to begin an exploration of ultrasound as an intersystemic facilitative adjunct to SPT for people with AOS. The results reported here represent early phases of clinical inquiry on this topic. For this participant it appeared that ultrasound visual feedback could be

used as a component in a structured SPT treatment approach to improve the production of treated and some untreated speech sounds. SPT alone appeared to be equally as impactful for improving speech treated speech sounds. The outcomes of the analysis appear to point to certain untreated speech sounds showing greater improvements when ultrasound was included in SPT. However, the limitations in the methodology of this study prevent definitive conclusions from these outcomes. A schema perspective of the results of this research could be that by providing visual biofeedback with the somatosensory and acoustic feedback inherent to SPT, the motor plans for lingually-focused articulations became stronger, resulting in improved execution of lingual movement. From an even broader perspective of multi-sensory speech processing, it would appear that intersystemic treatment approaches for AOS may be more effective than a unitary approach that relies exclusively on the inherent acoustic feedback of speech.

In terms of clinical implications that can be drawn from this study, using ultrasound in SPT for an individual with AOS was observed to have various benefits and challenges. Primarily among these challenges was remaining faithful to the SPT protocol. A second challenge around using ultrasound in treatment for AOS was the length of time required for training. This unanticipated training time altered the originally intended treatment schedule. However, the visual biofeedback provided by ultrasound appeared to neatly complement the SPT hierarchical process of treatment. It is also important to note that this study provides some evidence for the idea that speech and language treatment can be beneficial for individuals who present with severe deficits a year or more post-stroke. Despite receiving a WAB-R AQ (Kertez, 2006) score denoting a severe language deficit, the participant in this study made gains in both speech production as well as writing throughout treatment. A final clinical implication from this

research is that speech treatment using SPT appears to be particularly sensitive to overgeneralization effects and this study lends support to the notion that multiple targets should be selected for treatment when using this treatment approach for individuals with AOS. In conclusion, given a client who is sufficiently motivated and appropriately trained to use the tool, ultrasound could be a beneficial adjunct to a treatment approach for AOS in a clinical setting.

References

- Adler-Bock, M. (2004). Visual feedback from ultrasound in remediation of persistence /u/ errors: Case studies of two adolescents. The University of British Columbia. Vancouver: School of Audiology and Speech Sciences.
- Bacsfalvi, P. (2010). Attaining the lingual components of /1/ with ultrasound for three adolescents with cochlear implants. *Canadian Journal of Speech-Language Pathology*, *34*(3), 206-217.
- Bacsfalvi, P., & Bernhardt, B. M. (2011). Long-term outcomes of speech therapy for seven adolescents with visual feedback technologies: Ultrasound and electropalatography. *Clinical Linguistics and Phonetics*, 25(11-12), 1034-1043.
- Bailey, D. J., Eatchel, K., & Wambaugh, J. L. (2015). Sound Production Treatment: Synthesis and quantification of outcomes. *American Journal of Speech-Language Pathology*, 24(4), 798-814.
- Ballard, K. J., Granier, J. P., & Robin, D. A. (2000). Understanding the nature of apraxia of speech: Theory, analysis, and treatment. *Aphasiology*, 14(10), 969-995.
- Ballard, K. J., Wambaugh, J. L., Duffy, J. R., Layfield, C., Maas, E., Mauszycki, S., & McNeil, M. R. (2015). Treatment for acquired apraxia of speech: A systematic review of intervention research between 2004 and 2012. *American Journal of Speech-Language Pathology*, 24, 316-337.
- Ballard, K., & Robin, D. (2007). Influence of continual biofeedback on jaw pursuit-tracking in healthy adults and adults with apraxia plus aphasia. *Journal of Motor Behavior, 39*(1), 19-28.

- Bernhardt, B. M., Gick, B., Bacsfalvi, P., & Adler-Bock, M. (2005). Ultrasound in speech therapy with adolescents and adults. *Clinical Linguistics and Phonetics*, 19(6-7), 605-617.
- Bernhardt, M., Stemberger, J., & Charest, M. (2010). Intervention for speech production in cheldren and adolescents: Models of speech production and therapy approaches. *Canadian Journal of Speech-Language Pathology and Audiology*, 34(3), 157-167.
- Bislick, L. P., Weir, P. C., Spencer, K., Kendall, D., & Yorkston, Y. M. (2012). Do principles of motor learning enhance retention and transfer of speech skills? A systematic review. *Aphasiology*, 26(5), 709-728.

Dabul, B. L., (1979). Apraxia Battery for Adults. Pro-Ed.

- Dell, G. S., & O'Seaghdha, P. G. (1992). Stages of lexical access in language production. Cognition, 42, 287-314.
- Derrick, D., & Gick, B. (2013). Aerotactile integration drom distal skin stimuli. *Multisensory Research, 26*, 405-416.
- Dollaghan, C. A. (2007). *The handbook for evidence-based practice in communication disorders*. Baltimore: Paul H. Brookes Pub.
- Dworkin, J. P., & Culatta, R. A. (1980). *Dworkin-Culatta oral mechanism examination*. Edgewood Press.
- Fadiga, L., Craighero, L., Buccino, G., & Rizzolatti, G. (2002). Speech listening specifically modulates the excitability of tongue muscles: A TMS study. *European Journal of Neuroscience*, 15(2), 399-402.
- Fawcett, S., Bacsfalvi, P., & Bernhardt, B. M. (2008). Ultrasound as visual feedback in speech therapy for /1/ with adults with down syndrome. *Down Syndrome Quarterly*, *10*(1), 4-12.

- Guenther, F. H. (2006). Cortical interactions underlying the production of speech sounds. *Journal of Communication Disorders*, *39*, 350 - 365.
- Hickok, G., Houde, J., & Rong, F. (2011). Sensorimotor integration in speech processing computational basis and neural organization. *Neuron*, 69, 407-422.
- Howard, S., & Varley, R. (1995). Using electropalatography to treat severe acquired apraxia of speech. *European Journal of Disorders of Communication, 30*, 246-255.
- Katz, W. F., McNeil, M. R., & Garst, D. M. (2010). Treating apraxia of speech (AOS) with EMA-supplied visual augmented feedback. *Aphasiology*, *24*(6-8), 826-837.
- Kertesz, A. (2006). Western Aphasia Battery Revised. Grune & Stratton.
- Knock, T. R., Ballard, K. J., Robin, D. A., & Schmidt, R. A. (2000). Influence of order of stimulus presentation on speech motor learning: A principled approach to treatment for apraxia of speech. *Aphasiology*, 14(5-6), 653 - 668.
- Levelt, W. J., & Wheeldon, L. (1994). Do speakers have access to a mental syllabary? *Cognition*, 50, 239-269.
- Levelt, W., Roelefs, A., & Meyer, A. (1999). A theory of lexical access in speech production. Behavioural and Brain Sciences, 22, 1-75.
- Maas, E. (2006). The nature and time course of motor programming in apraxia of speech.
 University of California, San Diego. San Diego: Retrieved from: https://escholarship.org/uc/item/3vb0z9bx.
- Maas, E., Mailend, M.-L., & Guenther, F. H. (2015). Feedforward and feedback control in apraxia of speech: Effects of noise masking on vowel production. *Journal of Speech, Language, and Hearing Research, 58*, 185-200.

- 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.
- McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*(5588), 746-748.
- McNeil, M. R., Katz, W. F., Fossett, T. D., Garst, D. M., Szuminsky, N. J., Carter, G., & Lim,
 K. Y. (2010). Effects of online augmented kinematic and perceptual feedback on
 treatment of speech movements in apraxia of speech. *Folia Phoniatrica et Logopaedica*, 62(3), 127-133.
- McNeil, M. R., Robin, D. A., & Schmidt, R. A. (2009). Apraxia of speech: Definition differentiation and treatment. In M. R. McNeil, *Clinical management of sensorimotor speech disorders (2nd edition)* (pp. 249-268). New York: Thieme.
- Modha, G., Bernhardt, B. M., Church, R., & Bacsfalvi, P. (2008). Case study using ultrasound to treat /1/. *Journal of Language & Communication Disorders*, *43*(3), 323-329.
- Ogar, J., Slama, H., Dronkers, N., Amici, S., & Gorno-Tempini, M. (2005). Apraxia of speech: An overview. *Neurocase*, *11*(6), 427-432.
- Peach, R. K. (2005). Acquired apraxia of speech: Features, accounts, and treatment. *Topics In Stroke Rehabilitation, 11*(1), 49-58.
- Peirce, J. W. (2009). Generating stimuli for neuroscience using PsychoPy. Fontiers in Neuroinformatics, 2(10), 1 8.
- Pierce, J. W. (2007). PsychoPy Psychophysics software in Python. *Journal of Neuroscience Methods*, 162(1-2), 8 - 13.

- Preston, J. L., & Leaman, M. (2014). Ultrasound visual feedback for acquired apraxia of speech: A case report. *Aphasiology*, *28*(3), 278-295.
- Rosenbek, J. C., & Jones, H. H. (2009). Principles of treatment for sensorimotor speech disorders. In M. R. McNeil, *Clinical Management of Sensorimotor Speech Disorder (2nd edition)* (pp. 269-289). New York: Thieme.
- Schmidt, R. A. (1975). A schema theory of discrete motor skill learning. *Psychological Review*, 82(4), 225-260.
- Schmidt, R., & Lee, T. (1999). *Motor control and learning: A behavioral Emphasis (3rd ed)*. Windsor: Human Kinetics.
- Seitz, R. J., Matyas, T. A., & Carey, L. M. (2008). Neural plasticity as a basis for motor learning and neurorehabilitation. *Brain Impairment*, *9*(2), 103-113.
- Shawker, T. H., & Sonies, B. C. (1984). Ultrasound biofeedback for speech training: Instrumentation and preliminary results. *Investigative Radiology*, *20*, 90 - 93.
- Stone, M. (2005). A guide to analysing tongue motion from ultrasound images. Clinical Linguistics & Phonetics, 19:6-7, 455-501.
- Tourville, J. A., & Guenther, F. H. (2011). The DIVA model: A neural history of speech acquisition and production. *Language and Cognitive Processes*, *26*(7), 952-981.
- Varley, R., & Whiteside, S. P. (2001). What is the underlying impairment in acquired apraxia of speech? *Aphasiology*, 15(1), 39-84.
- Wambaugh, J. L., & Mauszycki, S. C. (2008). The effects of rate control treatment on consonant production accuracy in mild apraxia of speech. *Aphasiology*, *22*(7-8), 906-920.
- Wambaugh, J. L., & Mauszycki, S. C. (2010). Sound Production Treatment: Application with severe apraxia of speech. *Aphasiology*, 24(6-8), 814 - 825.

- Wambaugh, J. L., & Nessler, C. (2004). Modification of sound production treatment for apraxia of speech: Acquisition and generalisation effects. *Aphasiology*, 18(5-7), 407-427.
- Wambaugh, J. L., & Shuster, L. I. (2008). The nature and management of neuromotor speech disorders accompanying aphasia. In R. Chapey, *Language Intervention Strategies and Related neurogenic communication disorders (fifth edition)* (pp. 1009 1042).
 Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins.
- Wambaugh, J. L., Duffy, J. R., McNeil, M. R., Robin, D. A., & Rogers, M. A. (2006)a.
 Treatment guidelines for acquired apraxia of speech: A synthesis and evaluation of evidence. *Journal of Medical Speech-Language Pathology*, 14(2), xv-xxxiii.
- Wambaugh, J. L., Duffy, J. R., McNeil, M. R., Robin, D. A., & Rogers, M. A. (2006)b. Treatment guidelines for acquired apraxia of speech: Treatment descriptions and recommendations. *Journal of Medical Speech-Laguage Pathology*, 14(2), xxxv-lxvii.
- Wambaugh, J. L., Kalinyak-Fliszar, M. M., West, J. E., & Doyle, P. J. (1998). Effects of treatment for sound errors in apraxia of speech and aphasia. *Journal of Speech, Language, and Hearing Research, 41*(4), 725-743.
- Wambaugh, J. L., Martinez, A. L., McNeil, M. R., & Rogers, M. A. (1999). Sound production treatment for apraxia of speech: overgeneralization and maintenance effects. *Aphasiology*, 13(9-11), 821-837.
- Wambaugh, J. L., West, J. E., & Doyle, P. J. (1998). Treatment of apraxia of speech: Effects of targeting sound groups. *Aphasiology*, 12(7-8), 731-743.
- Wambuagh, J. L., Nessler, C., Cameron, R., & Mauszycki, S. C. (2013). Treatment for acquired apraxia of speech: examination of treatment intensity and practice schedule. *American Journal of Speech-Language Pathology*, 22(1), 84 - 102.

- Watkins, K. E., Strafella, A. P., & Paus, T. (2003). Seeing and hearing speech excited the motor system insvolved in speech production. *Neuropsychologia*, *41*, 989-994.
- Zeigler, W. (2002). Psycholinguistic and motor theories of apraxia of speech. *Seminars in Speech and Language, 23*(4), 231 - 244.

Appendices



Appendix A Mid-Sagittal Ultrasound Image of the Tongue

Appendix B Probe Word List

Ga	Ghee
Saw	See
No	Knee
Caw	Кеу
Za	Zee
Da	D
Та	Теа
Shaw	She
Cha	Chi
Jaw	G
Raw	Re
Law	Lee
Ma	Ме
На	Не
Thaw	Theme

Appendix C Treatment Hierarchies

C.1 SPT Block Treatment Hierarchy

Step 1. Interventionist provides a verbal model of target word in a carrier phrase and requests repetition of just the target word.

a. For an accurate response, give feedback -

Ask for 5 repetitions* and proceed to next word

b. For an inaccurate response, give feedback and -

And move to Step 2 of Hierarchy

Step 2. Interventionist shows the target word printed in orthography pointing out the sounds that were produced incorrectly then models the target word (no carrier phrase) and asks for repetition

- Do not offer Articulatory cues at this point

a. For an accurate response, give feedback -

Ask for 5 repetitions* and proceed to next word

b. For an inaccurate response, give feedback -

And move to Step 3 of Hierarchy

Step 3. Integral stimulation "Watch, listen, and say it with me" (Rather than "watch me, listen to me" we use the phrase "watch" and "listen" so that this step can include the introduction of the ultrasound during the next block of treatment. During this block when we use the phrase "watch, listen.." direct the participant's attention to the interventionists face.

a. For an accurate response, give feedback -

Ask for 5 repetitions* and proceed to next word

b. For an inaccurate response, give feedback -

And move to Step 4 of Hierarchy

Step 4. Articulatory placement cue: describe the necessary movement of articulators to achieve a

production of the target sounds and produce the target word in choral production

a. For an accurate response, give feedback -

Ask for 5 repetitions* and proceed to next word

b. For an inaccurate response, give feedback -

Proceed to next word

*feedback is provided for accuracy for approximately 3/5 productions

C.2 SPT Plus U/S Block Treatment Hierarchy

Step 1. Interventionist provides a verbal model of target word in a carrier phrase and requests repetition of just the target word.

a. For an accurate response, give feedback -

Ask for 5 repetitions* and proceed to next word

b. For an inaccurate response, give feedback and -

And move to Step 2 of Hierarchy

Step 2. Interventionist shows the target word printed in orthography pointing out the sounds that were produced incorrectly and shows clay tongue model in correct articulatory posture then models the target word (no carrier phrase) and asks for repetition. No articulatory placement cues are provided.

a. For an accurate response, give feedback -

Ask for 5 repetitions* and proceed to next word
b. For an inaccurate response, give feedback -

And move to Step 3 of Hierarchy

Step 3. Integral stimulation with ultrasound model: situate the probe under own chin and direct attention to the computer monitor while cueing to "watch, listen, and say it with me". Say the word in choral production while watching the movement of the tongue on the screen.

a. For an accurate response, give feedback -

Ask for 5 repetitions*^ and proceed to next word

b. For an inaccurate response, give feedback -

And move to Step 4 of Hierarchy

Step 4. Articulatory placement cue with the ultrasound: situate the probe under the clients chin and direct attention to the ultrasound image and describe the necessary movement of articulators to achieve a production of the target sounds and produce the target word in choral production.

a. For an accurate response, give feedback -

Ask for 5 repetitions*^ and proceed to next word

b. For an inaccurate response, give feedback -

Proceed to next word

*feedback is provided for accuracy for approximately 3/5 productions

^Repetitions are to be performed without the assistance of ultrasound

Appendix D Treatment Fidelity Checklists

D.1 SPT Block Treatment Fidelity Checklist

Directions: Listen to the audio recordings of treatment session. Follow this checklist 1- The word is modeled by interventionist in a carrier phrase - The word should be used first in a phrase that is spoken in its entirety by the interventionist. The client should then make an attempt to repeat only the last word in the phrase.

2 - Verbal feedback is provided about whether the client's production is correct - Verbal feedback can refer to either the correctness of a production or the incorrectness of a production as long as some verbal instruction can be heard to indicate whether the next step of the hierarchy will be applied or if the client will now attempt 5 repetitions of the word just produced. Verbal feedback indicating a correct production may include but is not limited to the following: "that's right", "good", "mhmm", "nice", "yeah". Verbal feedback indicating an incorrect production may include but is not limited to the following: "not quite", "almost", "let's try it again", "mm" (with falling intonation).

3 - An orthographic cue is provided by the interventionist - The interventionist must refer to the visual nature of the orthographic cue in order to be checked off on the score sheet. This can include making a reference to looking at, showing, or seeing the word to the client. The cue can also refer to the orthographic name of one or more letter in the word.

4 - Verbal feedback is provided about whether the client's production is correct following orthographic cue - Verbal feedback can refer to either the correctness of a production or the incorrectness of a production as long as some verbal instruction can be heard to indicate whether the next step of the hierarchy will be applied or if the client will now attempt 5 repetitions of the word just produced. Verbal feedback indicating a correct production may include but is not limited to the following: "that's right", "good", "mhmm", "nice", "yeah".
Verbal feedback indicating an incorrect production may include but is not limited to the following: "not quite", "almost", "let's try it again", "mm" (with falling intonation).
5 - Interventionist directs client's attention to "watch, listen, and say it with me" - Minor variations on this phrase are allowed including, "say it together", "try it with me", "watch and listen".

6 - Attempts are made to produce the word in choral production with the client - Interventionist and client production of word must overlap in order to be checked off on the score sheet. More than one attempt at achieving choral production can be made.

7 - Verbal feedback is provided about whether the client's production is correct - Verbal feedback can refer to either the correctness of a production or the incorrectness of a production as long as some verbal instruction can be heard to indicate whether the next step of the hierarchy will be applied or if the client will now attempt 5 repetitions of the word just produced. Verbal feedback indicating a correct production may include but is not limited to the following: "that's right", "good", "mhmm", "nice", "yeah". Verbal feedback indicating an incorrect production may include but is not limited to the following: "not quite", "almost", "let's try it again", "mm" (with falling intonation).

8 - Interventionist describes placement of oral structures to provide an articulatory cue -Articulatory cues may refer to any oral structure and make reference to movement of structures, contact between structures, relative force of contact, resulting sound produced, or tactile expectation of placement.

9 - Attempts are made to produce the word in choral production with the client - Interventionist and client production of word must overlap in order to be checked off on the score sheet. More than one attempt at achieving choral production can be made.

10 - Verbal feedback is provided about whether the client's production is correct - Verbal feedback can refer to either the correctness of a production or the incorrectness of a production as long as some verbal instruction can be heard to indicate whether the next step of the hierarchy will be applied or if the client will now attempt 5 repetitions of the word just produced. Verbal feedback indicating a correct production may include but is not limited to the following: "that's right", "good", "mhmm", "nice", "yeah". Verbal feedback indicating an incorrect production may include but is not limited to the following: "not quite", "almost", "let's try it again", "mm" (with falling intonation).

11 - Five repetitions of the word are attempted by the client - Repetitions do not need to all be correct.

12 - Verbal feedback is provided about whether the client's productions are correct at least one time - Verbal feedback can refer to either the correctness of a production or the incorrectness of a production as long as some verbal instruction can be heard to indicate whether the next step of the hierarchy will be applied or if the client will now attempt 5 repetitions of the word just produced. Verbal feedback indicating a correct production may include but is not limited to the following: "that's right", "good", "mhmm", "nice", "yeah". Verbal feedback indicating an incorrect production may include but is not limited to the following: "not quite", "almost", "let's try it again", "mm" (with falling intonation). Feedback can be provided during or after the five client attempts. If the feedback is provided after the last repetition attempts, the feedback must refer to more than one of the productions just attempted. Feedback of this nature

may include but is not limited to the following: "all except the last one was correct", "you were able to fix it when it started to sound off", "I heard three good ones but two were not quite right".

D.2 SPT Plus U/S Block Treatment Fidelity Checklist

Directions: Listen to the audio recordings of treatment session. Follow this checklist 1- The word is modeled by interventionist in a carrier phrase - The word should be used first in a phrase that is spoken in its entirety by the interventionist. The client should then make an attempt to repeat only the last word in the phrase.

2 - Verbal feedback is provided about whether the client's production is correct - Verbal feedback can refer to either the correctness of a production or the incorrectness of a production as long as some verbal instruction can be heard to indicate whether the next step of the hierarchy will be applied or if the client will now attempt 5 repetitions of the word just produced. Verbal feedback indicating a correct production may include but is not limited to the following: "that's right", "good", "mhmm", "nice", "yeah". Verbal feedback indicating an incorrect production may include but is not limited to the following: "not quite", "almost", "let's try it again", "mm" (with falling intonation).

3 - An orthographic cue is provided by the interventionist and may include reference to viewing a tongue model- The interventionist must refer to the visual nature of the cue in order to be checked off on the score sheet. This can include making a reference to looking at, showing, or seeing the word or tongue model to the client. The cue can also refer to the orthographic name of one or more letter in the word.

4 - Verbal feedback is provided about whether the client's production is correct following orthographic cue - Verbal feedback can refer to either the correctness of a production or the incorrectness of a production as long as some verbal instruction can be heard to indicate whether the next step of the hierarchy will be applied or if the client will now attempt 5 repetitions of the word just produced. Verbal feedback indicating a correct production may include but is not limited to the following: "that's right", "good", "mhmm", "nice", "yeah". Verbal feedback indicating an incorrect production may include but is not limited to the following: "that's right", "mm" (with falling intonation).
5 – Interventionist directs clients attention to ultrasound screen "watch me, listen, and say it with me", "watch and listen" as long as the cue directs the client to view the interventionist's production using the ultrasound.

6 - Attempts are made to produce the word in choral production with the client - Interventionist and client production of word must overlap in order to be checked off on the score sheet. More than one attempt at achieving choral production can be made.

7 - Verbal feedback is provided about whether the client's production is correct - Verbal feedback can refer to either the correctness of a production or the incorrectness of a production as long as some verbal instruction can be heard to indicate whether the next step of the hierarchy will be applied or if the client will now attempt 5 repetitions of the word just produced. Verbal feedback indicating a correct production may include but is not limited to the following: "that's right", "good", "mhmm", "nice", "yeah". Verbal feedback indicating an incorrect production may include but is not limited to the following: "not quite", "almost", "let's try it again", "mm" (with falling intonation).

8 – Interventionist directs attention to the ultrasound once again and makes reference to seeing the participant's articulation of the target words. You should hear the ultrasound turn on.
9 - Interventionist describes placement of the client's oral structures to provide an articulatory cue - Articulatory cues may refer to any oral structure and make reference to matching tongue shape to previously observed productions, intended or observed movement of structures, contact between structures, relative force of contact, resulting sound produced, or tactile expectation of placement.

10 - Verbal feedback is provided about whether the client's production is correct - Verbal feedback can refer to either the correctness of a production or the incorrectness of a production as long as some verbal instruction can be heard to indicate whether the next step of the hierarchy will be applied or if the client will now attempt 5 repetitions of the word just produced. Verbal feedback indicating a correct production may include but is not limited to the following: "that's right", "good", "mhmm", "nice", "yeah". Verbal feedback indicating an incorrect production may include but is not limited to the following: "not quite", "almost", "let's try it again", "mm" (with falling intonation).

11 - Five repetitions of the word are attempted by the client - Repetitions do not need to all be correct.

12 - Verbal feedback is provided about whether the client's productions are correct at least one time - Verbal feedback can refer to either the correctness of a production or the incorrectness of a production as long as some verbal instruction can be heard to indicate whether the next step of the hierarchy will be applied or if the client will now attempt 5 repetitions of the word just produced. Verbal feedback indicating a correct production may include but is not limited to the following: "that's right", "good", "mhmm", "nice", "yeah". Verbal feedback indicating an

incorrect production may include but is not limited to the following: "not quite", "almost", "let's try it again", "mm" (with falling intonation). Feedback can be provided during or after the five client attempts. If the feedback is provided after the last repetition attempts, the feedback must refer to more than one of the productions just attempted. Feedback of this nature may include but is not limited to the following: "all except the last one was correct", "you were able to fix it when it started to sound off", "I heard three good ones but two were not quite right".

Appendix E Probe Word Elicitation Transcriptions

E.1 Baseline

	1	2	3	4	5
Ga	la	ka	ka	ka	kha
Ghee	sit	si	si	si	si
Saw	θal	ta	so		
See	hid	tid	sit	sit	sit
No	lod	not	no		
Knee	hid	sid	tid	lit	tid
Caw	tal	tid	ka		
Key	si	si	hid	zid	hid
Za	sa	zɛtə	szat	szad	zad
Zee	3i	zi			
Da	sa	sa	sa	ta	sa
D	dit	tit	dit	tid	di:d
Та	tai	te	hai	te	ha
Tea	əhit∫	nit	til	ti	
Shaw	ſɔ				
She	si	∫i			
Cha	фa	tʃɔ			
Chi	did	∫i	til	sit	sit
Jaw	∫sa	фar	∫ar	фаr	фаr
G	∫i	s∫i	∫i	∫ʒi	фi
Raw	ra				
Re	hi	rid	ri		
Law	hala	la			
Lee	θzi	si	li		
Ma	vad	pa	ma		
Me	zi	mi			
На	θod	θο			
He	sid	bid	bid	pid	vid
Thaw	la	ka	ka	ka	kha
Theme	sit	si	si	si	si

E.2 Pre-SPT

	1	2	3	4	5
Ga	hər	hor	si	∫3 [°]	∫a
Ghee	zi	si	si	si	hzi
Saw	sa				
See	si	si	si		
No	so	∫3 [°]	no		
Knee	si	si	ti	li	ri
Caw	∫ar	ha	ha	har	o:
Key	si	si	si	si	ki
Za	ha	sar	zar	zar	za
Zee	si	zi			
Da	si	ha	tar	ta	ta
D	i	si	i	zi	si
Та	θа	ha	tar	ta	
Tea	si	si	hli	si	si
Shaw	sar	sar	sar	sa	sar
She	∫i				
Cha	∫ar	sar	∫ra	har	∫ha
Chi	si	si	s∫i	∫i	si
Jaw	sə	sar	sa	sia	sa
G	zi	dzi	si	si	szi
Raw	ra				
Re	rar	ra	ri	ri	ri
Law	so	har	lo		
Lee	si	∫3 [⊾]	hli	li	li
Ma	ra	ma			
Me	si	iф	vi	si	mi
На	ha				
Не	zi	hi			
Thaw	θο				
Theme	siə	peə	θid	lid	lid

E.3 Post-SPT

	1	2	3	4	5
Ga	kar	kar	kar	kar	kar
Ghee	gi				
Saw	səl	so			
See	si				
No	go	go	no		
Knee	si	hli	sei	nin	sə
Caw	ko				
Key	zil	ki			
Za	za				
Zee	zi				
Da	sar	kar	kar	adedede	kar
D	gi	hli	ka	gi	gi
Та	ta				
Теа	si	si	tsi	tsi	si
Shaw	sar	sar	sar	∫ar	∫a
She	si	si	∫si	∫i	
Cha	sar	sar	sar	sar	sar
Chi	zi	szi	si	si	se
Jaw	sar	zar	∫⊃	saha	sar
G	zi	zi	si	həlo	3i
Raw	ra				
Re	ri				
Law	15				
Lee	θli	hli	li		
Ma	bar	ma			
Me	mi				
На	ha				
Не	hi				
Thaw	θο				
Theme	hibə	hip	θis	lib	fim

E.4 Pre-U/S

	1	2	3	4	5
Ga	ka	har	gar	gar	gar
Ghee	si	si	li	si	sīt
Saw	sar	sar	sar	sa~	sa
See	si				
No	lo	wor	∫or	no	
Knee	si	Ιου	loustəd	leid	hleɪd
Caw	har	har	kar	har	har
Key	si	si	hli	hi	hi
Za	sa	sar	sar	szar	sar
Zee	si	si	si	sli	zi
Da	si	har	har	har	har
D	si	si	si	si	si
Та	sar	sar	sei	tar	sar
Tea	si	si	∫i	∫ri	li
Shaw	sar	∫art	sar	∫ar	sar
She	si	si	si	s∫i	∫i
Cha	har	har	∫⊃	har	har
Chi	si	s∫i	si	si	si
Jaw	ſɔ	∫ar	∫ar	∫ar	∫ar
G	zi	zi	lo [:] rt	si	helə
Raw	ra	ra			
Re	ri				
Law	S3 [°]	har	har	lar	lo
Lee	lıŋ	hlıl	hə	sar	si
Ma	far	mar	mar	par	ma
Me	mi				
На	ha	ha			
Не	hi				
Thaw	sar	har	har	har	har
Theme	rid~	far~	towatə	per	θim

E.5 Post-U/S

	1	2	3	4	5
Ga	ga	ga	k∫ag	s sarg	
Ghee	gi				
Saw	so				
See	si				
No	no				
Knee	ni				
Caw	kə	k∫a	kag	k∫a	
Key	hi	ki			
Za	za				
Zee	zi				
Da	nta	nta	da		
D	ni	ndi	ndi	nid	Э
Та	sta	ta			
Tea	ti				
Shaw	∫ag	saris	∫⊃g	sag	sag
She	∫i				
Cha	so	so	sa	∫ar	∫ 3 [•]
Chi	tfi				
Jaw	ශ්ා				
G	க குi				
Raw	rah	rə			
Re	ri				
Law	15				
Lee	ji	li	f li		
Ma	ma				
Me	mi				
На	ha				
Не	hil	hin	hi		
Thaw	sak	sag	sak	∫agദ	gʒlə
Theme	sifə	fitæ	pīgol	pīgə	pil